



北京大学

本科生毕业论文

题目：随机模型中的相变与临界性

Phase Transitions and Criticality in Stochastic Models

姓 名：杜航

学 号：1900010611

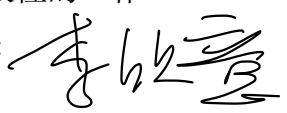
院 系：数学科学学院

专 业：数学与应用数学

导师姓名：李欣意

二〇二三 年 五 月

北京大学本科毕业论文导师评阅表

学生姓名	杜航	学生学号	1900010611	论文成绩	优
学院(系)	北京大学数学科学学院			学生所在专业	数学与应用数学
导师姓名	李欣意	导师单位/ 所在研究所	北京大学国际 数学研究中心	导师职称	助理教授
论文题目 (中、英文)		随机模型中的相变与临界性 Phase transitions and criticality in stochastic models			
<p style="text-align: center;">导师评语</p> <p>(包含对论文的性质、难度、分量、综合训练等是否符合培养目标的目的等评价)</p> <p>此篇论文聚焦于相变与临界性这两个概率论与理论计算机和统计物理的交叉领域中受到广泛关注的研究话题，对于不同模型的多个基本问题展开了深入的研究。在相关随机图模型的相关性检验与匹配复原问题上，该论文给出了信息论相变点的精确刻画，在信息论意义下回答了此模型的两个基本问题；在独立随机图的图匹配问题中，该论文给出了严格的信息论阈值与算法阈值，并以此建立了有效算法意义上的相变，为图匹配问题在随机环境下的性态提供了新角度的理解；在电缆图上的布朗圈汤模型中，该论文给出了一大类电缆图上渗流临界点位置，完善了该模型渗流理论的图景；在临界态平面伯努利渗流模型中，该论文考察了扮演着重要角色的臂事件，并对其中一大类事件概率给出精确渐进估计，大幅改进了前人结果。</p> <p>本篇论文研究范围广泛，选题角度着眼于概率论与理论计算机及统计物理中自然而基本的问题，研究方法涉及多领域的知识与技术，最终结果创新性相当高，达到了国际一流学术杂志的发表水平。论文部分工作已将见刊于信息论领域顶级杂志 IEEE Transanction on Information Thoery 及统计学顶级杂志 Annals of Statistics, 其余部分工作也处于审稿流程中。从各方面讲，本论文都远远超过本科生毕业论文水准，并可以证明作者已经具备了基本的数学学术素养与科研能力，这与数学科学学院的本科生培养目标是高度契合的。</p> <p>综上，我认为这是一篇质量极高的本科生毕业论文。该论文充分展现了作者在本科期间所培养出的科研能力以及令人惊叹的学术活力。希望作者在数学研究的道路上不忘初心，砥砺前行，争取做出更加深刻的、具有突破性的工作。</p> <p style="text-align: right;">导师签名: </p> <p style="text-align: right;">2023 年 5 月 30日</p>					

版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则一旦引起有碍作者著作权之问题，将可能承担法律责任。

摘要

该论文是对作者在本科学习期间所做科研工作的一个汇总。这些工作主要关注随机模型中的相变现象与临界态性质，系作者本人与常寅山（四川大学），丁剑（北京大学），高一帆（香港城市大学），巩舒阳（北京大学），黄润东（北京大学），李欣意（北京大学）和庄子杰（美国宾夕法尼亚大学）共同合作完成。这一系列的工作着眼于来自概率论的理论视角与技术手段在理解理论计算机与统计物理中诸多有趣现象中所扮演的角色。该论文探究了不同随机模型中的相变与临界性问题，其中相变现象涵盖信息论相变，算法相变以及动力学相变等多个方面。该论文的研究范围涉及到相关随机图模型的相关性检验与匹配复原，独立随机图的图匹配，电缆图上布朗圈汤的渗流相变点以及临界态平面伯努利渗流的臂事件概率等诸多具体问题。

关键词：相变, 临界性, 随机图, 渗流

Phase Transitions and Criticality in Stochastic Models

Hang Du (Mathematics and applied mathematics)

Directed by Prof. Xinyi Li

ABSTRACT

This thesis comprises a collection of research works conducted by the author during the undergraduate study, with a focus on phase transitions and criticality in stochastic models, collaborating with Yinshan Chang (Sichuan University), Jian Ding (Peking University), Yifan Gao (City University, Hong Kong), Shuyang Gong (Peking University), Rundong Huang (Peking University), Xinyi Li (Peking University) and Zijie Zhuang (University of Pennsylvania, U.S.A.). The studies emphasize probabilistic perspectives and approaches to understanding intriguing phenomena in theoretical computer science and statistical physics. The thesis explores informational phase transitions, computational phase transitions, dynamical phase transitions, and phenomena at criticality in various models and problems. The scope of the thesis includes detection and matching recovery for correlated random graphs, graph alignment problem over independent random instances, percolation threshold for Brownian loop soup on cable graphs, and arm events in critical planar Bernoulli percolation.

KEY WORDS: phase transition, criticality, random graph, percolation

Contents

1	Overview	5
2	Informational phase transition of correlation detection between a pair of random graphs	7
2.1	Introduction	7
2.2	Detect correlation via densest subgraph	12
2.2.1	The densest subgraph of an Erdős-Rényi graph	12
2.2.2	Test graph correlation	14
2.3	Impossibility for detection	15
2.3.1	The conditional second moment method	16
2.3.2	Truncation for the π^* -intersection graph	21
2.3.3	The truncated exponential moment	25
3	Informational phase transition of partial recovery in the correlated random graph model	34
3.1	Introduction	34
3.1.1	Background and related results	35
3.1.2	Connection to previous works	37
3.1.3	Notations	38
3.2	Possibility for partial recovery	40
3.2.1	Orbits and tail probabilities	42
3.2.2	Proof of Proposition 3.4	45
3.3	Impossibility for partial recovery	48
3.3.1	Truncations for the second moment	50
3.3.2	Proof of Proposition 3.11 via second moment bound	54
3.4	Complimentary proofs	57
3.4.1	Proof of Proposition 3.6	57

3.4.2	Proof of Lemma 3.9	59
3.4.3	Proof of Proposition 3.15	60
4	A Polynomial-time approximation scheme for the maximal overlap of two independent Erdős-Rényi graphs	77
4.1	Introduction	77
4.1.1	Background and related results	79
4.1.2	An overview of our method	80
4.2	Proof of Theorem 4.1	83
4.2.1	Upper bound of the maximal overlap	83
4.2.2	Preliminaries of the algorithm	83
4.2.3	Description of the algorithm	89
4.2.4	Aanalysis of the algorithm	92
4.3	Complimentary proofs	95
4.3.1	Proof of Proposition 4.8	96
4.3.2	Proof of Proposition 4.9	102
4.3.3	Proof of Proposition 4.10	106
5	Algorithmic phase transition of the random graph alignment problem	121
5.1	Introduction and main results	121
5.1.1	Backgrounds and prior works	126
5.1.2	Proof overview	128
5.2	Proof of informational results	130
5.2.1	Informational upper bounds	130
5.2.2	Informational lower bounds for $p \gg p_c$	131
5.3	Algorithmic lower bounds via greedy matching	138
5.3.1	The algorithm framework and proof outlines	138
5.3.2	Analysis for sparse regime	142

5.3.3	Analysis for dense regime	147
5.4	Computational hardness for online algorithms	150
5.4.1	The forbidden structure	151
5.4.2	Completing the proof	154
5.5	Complimentary proofs	156
5.5.1	Tail Bounds	156
5.5.2	Proof of Lemma 5.19	158
5.5.3	Proof of Proposition 5.14	160
5.5.4	Proof of Proposition 5.20	160
6	Percolation phase transition of Brownian loop soup on metric graphs	172
6.1	Introduction and the main result	172
6.2	Preliminaries	173
6.3	Proof of the main result	176
7	Sharp asymptotics for arm probabilities in critical planar percolation	182
7.1	Introduction	182
7.1.1	Main results	184
7.1.2	Comments	186
7.2	Notation and preparatory results	190
7.3	Notation and conventions	190
7.3.1	Setting for percolation	191
7.3.2	Some classical tools	193
7.3.3	Arm events and asymptotics	194
7.3.4	Faces	198
7.3.5	Separation lemmas in the half-plane	200
7.3.6	Separation lemmas in the plane	205
7.3.7	Percolation exploration process and scaling limits	207

7.4	Couplings and conditional arm probabilities	211
7.4.1	The half-plane case	211
7.4.2	The plane cases	216
7.5	Proof of main theorems	219
7.5.1	Comparison estimates in the half-plane	219
7.5.2	Proof of Theorem 7.1	225
7.5.3	Comparison estimates in the plane	226
7.5.4	Proof of Theorems 7.2 and 7.3	232
7.6	Proof of coupling results	233
7.6.1	$j \geq 2$ arms in the half-plane	233
7.6.2	Coupling arm events in the plane	239
7.7	Complimentary proofs	242
7.7.1	Proof of Lemmas 7.11 and 7.13	242
7.7.2	Proof of Lemma 7.14	243
7.7.3	Proof of the strong separation lemma	244
	Acknowledgements	249

1 Overview

Phase transition is a concept from statistical physics that refers to the significant changes within a system when its controlling parameters surpass certain thresholds. Likewise, the study of criticality, also derived from physics, is concerned with examining the unique characteristics of a system at these exact critical points. With the development of modern mathematical physics, perspectives and ideas from physics have contributed to notable insights in numerous fields of pure and applied mathematics. Specifically, the terms phase transition and criticality have evolved to acquire new meanings within the contexts of probability theory and stochastic processes.

Modern probability theory exists at the intersection point of mathematics and statistical physics, and offers a vast array of probabilistic models that aim to comprehend the nature of the real world. Simultaneously, these models provide a rigorous perspective to aid the development of physical theories. Occurrences of phase transitions are pretty common within these stochastic models, and understanding the transitional thresholds as well as the behavior of models at critical points has been a consistent topic of interest in this field.

This thesis also focuses on various phase transitions and characteristics at the critical point of several distinct stochastic models. These models include the correlated random graph model, the graph alignment problem over random instances, the Brownian loop soup percolation model on cable graphs, and the planar Bernoulli percolation model. In Sections 2 and 3, we explore the informational phase transitions related to two essential questions concerning the correlated random graph model, namely, correlation detection and matching recovery. These sections are based on two joint works arXiv:2203.14573 and arXiv:2205.14650 with Jian Ding. Sections 4 and 5 address the graph alignment problem for two independent Erdős-Rényi graphs, examining both the informational and algorithmic phase transitions. These sections are based on a joint work arXiv:2210.07823 with Jian Ding and Shuyang Gong together with an ongoing joint work with Shuyang Gong and Rundong Huang. In Section 6, we establish the percolation threshold for the Brownian loop soup

on an extensive class of cable graphs. This completes the picture of the dynamical phase transition for this model. This section is based on a joint work [arXiv:2304.08225](#) with Xinyi Li and Yinshan Chang. Lastly, in [Section 7](#), we scrutinize arm events in the critical planar percolation. By establishing sharp asymptotics for their probabilities, we reveal a fascinating property unique to the critical point. This section is based on a joint work [arXiv:2205.15901](#) with Yifan Gao, Xinyi Li and Zijie Zhuang.

Each section of this thesis is self-contained and can be read independently. Notations are applied separately within each section for the ease of understanding.

2 Informational phase transition of correlation detection between a pair of random graphs

The problem of detecting edge correlation between two Erdős-Rényi random graphs on n unlabeled nodes can be formulated as a hypothesis testing problem: under the null hypothesis, the two graphs are sampled independently; under the alternative, the two graphs are independently sub-sampled from a parent graph which is Erdős-Rényi $\mathbf{G}(n, p)$ (so that their marginal distributions are the same as the null). We establish a sharp information-theoretic threshold when $p = n^{-\alpha+o(1)}$ for $\alpha \in (0, 1]$ which sharpens a constant factor in a recent work by Wu, Xu and Yu. A key novelty in our work is an interesting connection between the detection problem and the densest subgraph of an Erdős-Rényi graph. This section is based on a joint work with Jian Ding.

2.1 Introduction

In this section, we study the information-theoretic threshold for detecting the correlation between a pair of Erdős-Rényi graphs. To put this question into a precise mathematical framework, we first need to choose a probabilistic model for a pair of correlated Erdős-Rényi graphs, and one natural choice is to obtain the two graphs as two independent subsamplings from a common Erdős-Rényi graph. More formally, for two vertex sets V, \mathbf{V} of cardinality n we let E_0 be the set of unordered pairs (u, v) with $u, v \in V$ and $u \neq v$, and let \mathbf{E}_0 be the set of unordered pairs (\mathbf{u}, \mathbf{v}) with $\mathbf{u}, \mathbf{v} \in \mathbf{V}$ and $\mathbf{u} \neq \mathbf{v}$. For some model parameters $p, s \in (0, 1)$ (which may depend on n), we sample a uniform bijection π^* between V and \mathbf{V} , independent Bernoulli variables $\{J_{u,v} : (u, v) \in E_0\}$ with parameter p and independent Bernoulli variables $\{I_{u,v} : (u, v) \in E_0\}, \{\mathbf{l}_{\mathbf{u}, \mathbf{v}} : (\mathbf{u}, \mathbf{v}) \in \mathbf{E}_0\}$ with parameter s , and then we define

$$(G_{u,v}, \mathbf{G}_{\pi^*(u), \pi^*(v)}) = (J_{u,v} I_{u,v}, J_{u,v} \mathbf{l}_{\pi^*(u), \pi^*(v)}). \quad (2.1)$$

Then (G, \mathbf{G}) forms a pair of correlated Erdős-Rényi graphs where the edge set of E (respectively \mathbf{E}) consists of all $(u, v) \in E_0$ (respectively $(u, v) \in \mathbf{E}_0$) such that $G_{u,v} = 1$ (respectively $\mathbf{G}_{u,v} = 1$). Note that marginally each G and \mathbf{G} is an Erdős-Rényi graph on n vertices with edge probability ps , whose law we denote as $\mathbf{G}(n, ps)$.

Therefore, one (natural) version for the problem of detecting correlated Erdős-Rényi graphs can be formulated as a hypothesis testing problem, where the null hypothesis H_0 and the alternative hypothesis H_1 are given as

$$H_0 : (G, \mathbf{G}) \sim \mathbf{P}, \text{ which is a pair of independent Erdős-Rényi Graphs } \mathbf{G}(n, ps),$$

$$H_1 : (G, \mathbf{G}) \sim \mathbf{Q}, \text{ which is a pair of correlated graphs given by the rule (2.1).}$$

Our goal is to test H_0 versus H_1 given (G, \mathbf{G}) as observations while π^* remains to be unknown. It is well-known that the testing error is captured by the total variation distance between the null and alternative distributions, and our main contribution is to establish a sharp phase transition on this total variation distance in the sparse regime.

Theorem 2.1. *Suppose $p = p(n)$ satisfies $p = n^{-\alpha+o(1)}$ for some $\alpha \in (0, 1]$ as $n \rightarrow \infty$. Let $\lambda_* = \varrho^{-1}(\frac{1}{\alpha})$ (where ϱ is defined in (2.5) below), then for any constant $\varepsilon > 0$, the following holds. For $\alpha \in (0, 1]$, if s satisfies $nps^2 \geq \lambda_* + \varepsilon$, then*

$$\text{TV}(\mathbf{P}, \mathbf{Q}) = 1 - o(1) \text{ as } n \rightarrow \infty, \quad (2.2)$$

where $\text{TV}(\mathbf{P}, \mathbf{Q}) = \frac{1}{2} \sum_{\omega} |\mathbf{P}[\omega] - \mathbf{Q}[\omega]|$ is the total variation distance between \mathbf{P} and \mathbf{Q} . In addition, for $\alpha \in (0, 1)$, if s satisfies $nps^2 \leq \lambda_* - \varepsilon$, then

$$\text{TV}(\mathbf{P}, \mathbf{Q}) = o(1) \text{ as } n \rightarrow \infty. \quad (2.3)$$

Our work is closely related to and much inspired by a recent work [45], where a sharp threshold was established for $\alpha = 0$ and upper and lower bounds on λ_* up to a constant factor were established for $\alpha \in (0, 1]$. In particular, Theorem 2.1 solves [45, Section 6, Open Problem 3]. It is worth emphasizing that (2.3) for $\alpha = 1$ with $np \rightarrow \infty$ was already

proved in [45] (which even allows $\epsilon \rightarrow 0$ as long as $\epsilon \gg n^{-1/3}$). While our method should also be able to give (2.3) for $\alpha = 1$, we chose to exclude this case since the assumption $\alpha < 1$ allows to avoid some technical complications. Thus, the only remaining case is when np has order 1, in which there is no sharp phase transition. Indeed, on the one hand, for some constant $c > 0$ satisfying $np \geq c$ and $s \geq c$ we have

$$\liminf_{n \rightarrow \infty} \text{TV}(\mathbf{P}, \mathbf{Q}) > 0. \quad (2.4)$$

This follows readily by comparing the marginal distribution of the pair

$$\left(\frac{|E| - \binom{n}{2}ps}{\sqrt{\binom{n}{2}ps(1-ps)}}, \frac{|E| - \binom{n}{2}ps}{\sqrt{\binom{n}{2}ps(1-ps)}} \right)$$

under \mathbf{P} and \mathbf{Q} . A straightforward application of Central Limit Theorem yields that the marginal law of such pair is approximately a pair of independent normal variables with mean zero and variance 1 under \mathbf{P} , and in contrast is a pair of bivariate normal variables with mean zero, variance 1 and with correlation at least c under \mathbf{Q} . On the other hand, as shown in [45], if $s \leq 0.1$ and $n^{1/3}(1 - nps^2) \rightarrow \infty$ then $\limsup_{n \rightarrow \infty} \text{TV}(\mathbf{P}, \mathbf{Q}) < 1$.

Background and related results. Recently, there has been extensive study on the problem of detecting correlation between two random graphs and the closely related problem of matching the vertex correspondence in the presence of correlation. Questions of this type have been raised from various applied fields such as social network analysis [30, 31], computer vision [9, 3], computational biology [37, 38] and natural language processing [20].

Despite the fact that Erdős-Rényi Graph perhaps does not quite capture important features for any network arising from realistic problems, (similar to most problems on networks) it is plausible that a complete understanding for the case of Erdős-Rényi Graphs forms an important and necessary step toward the much more ambitious goal of mathematically understanding graph detection and matching problems for realistic networks arising from applications (note that for many applications it remains a substantial challenge to propose a reasonable underlying random graph model). Along this line, many progress has

been made recently, including information-theoretic analysis [8, 7, 19, 45, 44] and proposals for various efficient algorithms [36, 46, 25, 23, 17, 38, 3, 13, 5, 9, 10, 32, 16, 20, 15, 27, 28]. Out of these references, the ones closely related to our work include (the aforementioned) [45] and [44] which studied the information-theoretic threshold for the matching problem, as well as [28] which obtained an efficient algorithm for detection when the correlation between the two graphs is above a certain constant. As of now, a huge information-computation gap remains for both detection and matching problems, and it is a major challenge to completely understand the phase transition for the computational complexity for either detection or matching problems.

Recently, detection and matching problems have also been studied for models other than Erdős-Rényi. For instance, a model for correlated randomly growing graphs was studied in [35], graph matching for correlated stochastic block model was studied in [36] and graph matching for correlated random geometric graphs was studied in [40]. A very interesting direction is to design efficient algorithms for graph detection and matching that is robust to the underlying random graph models.

A connection to densest subgraph. In [45], the authors used the maximal overlap between the two graphs over all vertex bijections as the testing statistic. This is a natural and likely efficient statistic, although it is not so easy to analyze the maximal overlap in the correlated case so [45] lower-bounded it by the overlap given by the true matching. While this relaxation manages to capture the detection threshold in the dense regime (when $p = n^{o(1)}$), it only captures the threshold up to a constant factor in the sparse regime (when $p = n^{-\alpha+o(1)}$ for $\alpha > 0$). Here is a brief description on the insights behind that guided this paper: In the dense regime, near the threshold the intersection of two correlated random graphs is an Erdős-Rényi with large degree and thus it is more “regular” in a sense that the densest subgraph is more or less as dense as the whole graph. In the sparse regime, on the one hand, the intersection is an Erdős-Rényi with constant degree and in this case the spatial fluctuation plays a non-negligible role such that the densest subgraph has

significantly higher average degree than the whole graph; on the other hand, the maximal intersection of two independent random graphs (over all vertex bijections) is much more “regular” than an Erdős-Rényi such that its densest subgraph has about the same average degree as the whole graph (as proved in (2.2.2)). In summary, this suggests that a more efficient testing statistic is the maximal densest subgraph over all vertex bijections as defined in (2.9), which indeed yields the correct upper bound on the detection threshold (see Theorem 2.5). The major technical contribution of this paper is then to prove the sharp lower bound on the detection threshold (i.e., as in (2.3)). To this end, we use a similar truncated second moment method as employed in [45] with the additional insight that the truncation should be related to the densest subgraph. We will discuss more on this in Section 2.3.

Next we briefly describe the development on the densest subgraph for an Erdős-Rényi graph. This problem arose in the study of load balancing problem [25], a particular example of which is to balance the loads when assigning m balls into n bins subject to the constraint that each ball is assigned to either of two randomly chosen bins. The load balancing problem is also closely related to the emergence of a k -core in an Erdős-Rényi graph (a k -core is a maximal connected subgraph in which all vertices have degree at least k), and much progress has been made in [8, 20, 24, 21]. While [8, 20] made significant progress in understanding the densest subgraph (note that [24, 21] are in the context of hypergraphs), the asymptotic behavior for the maximal subgraph density of an Erdős-Rényi graph (with average degree of order 1) was only established in [3]. In fact, the authors of [3] managed to compute the asymptotic value for the maximal subgraph density of a random graph with prescribed degree sequence using the objective method from [2], and their result on Erdős-Rényi graphs (see Proposition 3.2) is crucial for our work.

Acknowledgements. We warmly thank Nicholas Wormald, Yihong Wu and Jiaming Xu for stimulating discussions. Hang Du is partially supported by the elite undergraduate training program of School of Mathematical Science in Peking University.

2.2 Detect correlation via densest subgraph

2.2.1 The densest subgraph of an Erdős-Rényi graph

The following result of [3] provides an important input for the proof of Theorem 2.1.

Proposition 2.2 ([3], Theorem 1, Theorem 3). *For any constant $\lambda > 0$, there exists a constant $\varrho(\lambda) > 0$ which can be explicitly written via a variational characterization, such that for an Erdős-Rényi graph $\mathcal{H}(V, \mathcal{E}) \sim \mathbf{G}(n, \frac{\lambda}{n})$,*

$$\max_{\emptyset \neq U \subset V} \frac{|\mathcal{E}(U)|}{|U|} \rightarrow \varrho(\lambda) \text{ in probability as } n \rightarrow \infty, \quad (2.5)$$

where $\mathcal{E}(U)$ is the collection of edges in \mathcal{E} with both endpoints in U . Further, the function $\varrho(\cdot)$ satisfies

$$1 \leq \frac{\varrho(\beta)}{\varrho(\alpha)} \leq \frac{\beta}{\alpha}, \quad \forall 0 < \alpha < \beta, \quad (2.6)$$

hence $\varrho(\cdot)$ is continuous and increasing.

Remark 2.3. Although [3] treated Erdős-Rényi graphs with n vertices and $\lfloor \lambda n \rfloor$ edges for a fixed constant $\lambda > 0$, counterparts of all results apply to the $\mathbf{G}(n, \frac{\lambda}{n})$ model since the total number of edges in $\mathbf{G}(n, \frac{\lambda}{n})$ concentrate around $\frac{\lambda n}{2}$. In addition, our definition of ϱ is different from that in [3] by a scaling factor 2, i.e. $\varrho(\lambda)$ for us equals to $\varrho(\frac{\lambda}{2})$ for ϱ in [3].

We call the largest subgraph that maximizes the left hand side of (2.5) (if there are many such subgraphs, pick one of them arbitrarily) as *the densest subgraph*. In order to establish a sharp threshold phenomenon for graph detection, we need ϱ to be a *strictly* increasing function, as proved in the following proposition.

Proposition 2.4. *For $\lambda > 1$, $\varrho(\lambda) > 1$ and ϱ is strictly increasing. Furthermore, there exists some constant $c_\lambda > 0$, such that with probability tending to 1 as $n \rightarrow \infty$, the size of the densest subgraph in an Erdős-Rényi graph $\mathcal{H} \sim \mathbf{G}(n, \frac{\lambda}{n})$ is at least $c_\lambda n$.*

It is readily to see that $\varrho(\lambda) = 1$ for $\lambda \leq 1$ (For $\lambda < 1$ this follows from the well-known fact [14] that with probability tending to 1, an Erdős-Rényi graph in the sub-critical phase

has a component of size of order $\log n$ and all components have at most one cycle; for $\lambda = 1$ this follows from continuity). By Proposition 2.4, we can define the inverse function $\varrho^{-1} : [1, \infty) \rightarrow [1, \infty)$, where we let $\varrho^{-1}(1) = 1$. In the proof of Proposition 2.4 and some estimates later, we need the following Chernoff bound for Bernoulli variables (see [29, Theorem 4.4]): For $X \sim \text{Bin}(n, p)$, denote $\mu = np$, then for any $\delta > 0$,

$$\mathbb{P}[X \geq (1 + \delta)\mu] \leq \exp(-\mu[(1 + \delta) \log(1 + \delta) - \delta]). \quad (2.7)$$

Proof of Proposition 2.4. First we show $\varrho(\lambda) > 1$ for any $\lambda > 1$. This follows immediately from the fact that with probability tending to 1, the 2-core in \mathcal{H} contains $(1 - x)(1 - \frac{x}{\lambda} + o(1))n$ vertices and $(1 - \frac{x}{\lambda} + o(1))^2 \frac{\lambda n}{2}$ edges, where $x \in (0, 1)$ is given by the equation $xe^{-x} = \lambda e^{-\lambda}$ (See [35, Theorem 3] and see also e.g. [22, Lemma 2.16]).

Next we show that with probability tending to 1, the number of vertices in the densest subgraph has at least $c_\lambda n$ vertices for some constant $c_\lambda > 0$. Since $\varrho > 1$, there exists a $\rho = \rho(\lambda) > 1$ such that with probability tending to 1 the densest subgraph has edge-vertex ratio greater than ρ . Let \mathcal{A} be the event that there exists a subgraph of \mathcal{H} which has at most $c_\lambda n$ vertices and has edge-vertex ratio greater than ρ . Then, by a simply union bound,

$$\begin{aligned} \mathbb{P}[\mathcal{A}] &\leq \sum_{k \leq c_\lambda n} \binom{n}{k} \mathbb{P}\left[\text{Bin}\left(\frac{k(k-1)}{2}, \frac{\lambda}{n}\right) \geq \rho k\right] \\ &\stackrel{(2.7)}{\leq} \sum_{k \leq c_\lambda n} \frac{n^k}{k!} \exp\left(-\rho k \log\left(\frac{2n}{\lambda(k-1)}\right) + \rho k\right) \\ &\leq \sum_{k \leq c_\lambda n} \exp\left(-(\rho - 1)k \log\left(\frac{n}{k}\right) + Ck\right), \end{aligned}$$

where $\text{Bin}(\frac{k(k-1)}{2}, \frac{\lambda}{n})$ (similar notation applies below later) denotes a binomial variable (which is the distribution for the number of edges in a subgraph with k vertices), and $C = C(\lambda, \rho)$ is a large constant. Once we choose $c_\lambda > 0$ small enough such that $(\rho - 1) \log c_\lambda^{-1} > C$, then $\mathbb{P}[\mathcal{A}] = o(1)$, yielding the desired result.

Now we are ready to show ϱ is strictly increasing by a simple sprinkling argument. For

any $1 < \lambda_1 < \lambda_2$, we have shown that the densest subgraph A in $\mathcal{H}_{\lambda_1} \sim \mathcal{G}\left(n, \frac{\lambda_1}{n}\right)$ has at least $c_{\lambda_1}n$ vertices with probability tending to 1. We now independently sample another Erdős-Rényi graph $\mathcal{H}_{\lambda_2-\lambda_1} \sim \mathcal{G}\left(n, \frac{\lambda_2-\lambda_1}{n}\right)$ in the same vertex set as for \mathcal{H}_{λ_1} , then with probability tending to 1, $\mathcal{H}_{\lambda_2-\lambda_1}$ contains more than $\frac{(\lambda_2-\lambda_1)c_{\lambda_1}^2}{4}n$ edges within A . Since \mathcal{H}_{λ_2} stochastically dominates $\mathcal{H}_{\lambda_1} \cup \mathcal{H}_{\lambda_2-\lambda_1}$, this gives $\varrho(\lambda_2) > \varrho(\lambda_1)$. \square

2.2.2 Test graph correlation

We next define our testing statistic. For any bijection $\pi : V \rightarrow V$, we define the π -intersection graph \mathcal{H}_π of G and G as

$$\mathcal{H}_\pi = (V, \mathcal{E}_\pi), \text{ where } (u, v) \in \mathcal{E}_\pi \text{ if and only if } G_{u,v} = G_{\pi(u),\pi(v)} = 1. \quad (2.8)$$

Then our testing statistic is defined by

$$\mathcal{T}(G, G) \triangleq \max_{\pi} \max_{U \subset V: |U| \geq n/\log n} \frac{|\mathcal{E}_\pi(U)|}{|U|}. \quad (2.9)$$

Theorem 2.5. *Suppose $p = p(n)$ satisfies $p = n^{-\alpha+o(1)}$ for some $\alpha \in (0, 1]$ as $n \rightarrow \infty$. Let $\lambda_* = \varrho^{-1}(\frac{1}{\alpha})$. For s satisfying $nps^2 \geq \lambda_* + \varepsilon$, let $\tau = \frac{\varrho(\lambda_*) + \varrho(\lambda_* + \varepsilon)}{2}$ and let $\mathcal{T}(G, G)$ be defined as in (2.9). Then $\mathcal{T}(G, G)$ with threshold τ achieves strong detection, i.e.*

$$\mathbb{Q}[\mathcal{T}(G, G) < \tau] + \mathbb{P}[\mathcal{T}(G, G) \geq \tau] = o(1). \quad (2.10)$$

Note that Theorem 2.5 implies (2.2).

Proof of Theorem 2.5. Denote $\lambda = nps^2$. By Proposition 2.4,

$$\frac{1}{\alpha} = \varrho(\lambda_*) < \tau < \varrho(\lambda_* + \varepsilon). \quad (2.11)$$

Under H_1 , we see that \mathcal{H}_{π^*} is an Erdős-Rényi graph $\mathcal{G}\left(n, \frac{\lambda}{n}\right)$. Since $\tau < \varrho(\lambda_* + \varepsilon)$, by Propositions 2.2 and 2.4, with probability $1 - o(1)$ the densest subgraph of \mathcal{H}_{π^*} has edge-vertex ratio at least τ and has at least $c_\lambda n \geq n/\log n$ many vertices. This implies that $\mathbb{Q}[\mathcal{T}(G, G) < \tau] = o(1)$.

Furthermore, by a union bound, we get that

$$\mathbb{P}[\mathcal{T}(G, \mathbb{G}) \geq \tau] \leq \sum_{k \geq n/\log n} \binom{n}{k}^2 k! \mathbb{P} \left[\text{Bin} \left(\frac{k(k-1)}{2}, (ps)^2 \right) \geq \tau k \right],$$

where $\binom{n}{k}^2$ accounts for the number of ways to choose k vertices $A \subset V$ and k vertices $\mathbf{A} \subset \mathbf{V}$, $k!$ accounts for the number of bijections between A and \mathbf{A} . Crucially, we only need to take a union bound over $k!$ bijections since the subgraph of \mathcal{H}_π on A , subject to the constraint that π maps A to \mathbf{A} , only depends on $\pi|_A$. Since $p = n^{-\alpha+o(1)}$, we have $\log(\frac{n}{kp}) \geq [\alpha + o(1)] \log n$. In addition, $\tau > \frac{1}{\alpha}$ as in (2.11). Then, we get that for a small positive constant $\delta = \delta(\alpha, \tau)$,

$$\begin{aligned} \mathbb{P}[\mathcal{T}(G, \mathbb{G}) \geq \tau] &\stackrel{(2.7)}{\leq} \sum_{k \geq n/\log n} \frac{n^{2k}}{k!} \exp \left(-\tau k \log \left(\frac{2\tau n}{\lambda kp} \right) + \tau k \right) \\ &= \sum_{k \geq n/\log n} \frac{n^{2k}}{k!} \exp \left(-(1 + 2\delta)k \log n + o(k \log n) \right). \end{aligned}$$

By a straightforward computation, we then get that when n is large enough,

$$\mathbb{P}[\mathcal{T}(G, \mathbb{G}) \geq \tau] \leq \sum_{k \geq n/\log n} \frac{n^{(1-\delta)k}}{k!} \leq \frac{n^{(1-\delta)\lfloor n/\log n \rfloor}}{\lfloor n/\log n \rfloor!} \sum_{t \geq 0} \left(\frac{n^{1-\delta}}{n/\log n} \right)^t = o(1).$$

This proves $\mathbb{Q}[\mathcal{T}(G, \mathbb{G}) < \tau] + \mathbb{P}[\mathcal{T}(G, \mathbb{G}) \geq \tau] = o(1)$ as required. \square

2.3 Impossibility for detection

This section is devoted to the proof of (2.3). The basic idea follows the framework of conditional second moment for the likelihood ratio as in [44], with the aforementioned additional key insight that the truncation should involve the densest subgraph. It turns out somewhat more convenient in our case to compute the conditional first moment for $\frac{d\mathbb{Q}}{d\mathbb{P}}$ when $(G, \mathbb{G}) \sim \mathbb{Q}$, which is equivalent to its second moment when $(G, \mathbb{G}) \sim \mathbb{P}$. We choose the first moment formulation since it is then convenient to consider truncation on the π^* -intersection graph \mathcal{H}_{π^*} sampled according to \mathbb{Q} (i.e., when the two graphs are correlated).

We learned from [44] that when computing the moments of the likelihood ratio, the concept of *edge orbits* (induced by a permutation) plays an important role. We streamline this intuition a little further and prove Lemma 2.6 in Section 2.3.1. In light of Lemma 2.6, it is natural to separate the edge orbits depending on whether they are entirely contained in the π^* -intersection graph \mathcal{H}_{π^*} (see (2.14)). With Lemma 2.6 at hand, the issue reduces to bounding the moment from edge orbits that are entirely in \mathcal{H}_{π^*} , which naturally calls for a truncation on \mathcal{H}_{π^*} . As a major difference between [44] and our work, instead of truncating \mathcal{H}_{π^*} as a pseudo forest as in [44], we truncate on the maximal subgraph density for \mathcal{H}_{π^*} and some other mild conditions on small subgraph counts in \mathcal{H}_{π^*} (see Section 2.3.2). In Section 2.3.3 we bound the truncated moment, where the truncation on the maximal subgraph density plays a crucial role since it rules out the possibility of creating a large number of edge orbits in \mathcal{H}_{π^*} by only fixing the values of the bijection on a small number of vertices.

2.3.1 The conditional second moment method

Let \mathcal{Q} be the probability measure on the sample space $\Omega = \{(G, \mathbb{G}, \pi^*)\}$ under H_1 , i.e. under \mathcal{Q} , π^* is a uniform bijection and conditioned on π^* , (G, \mathbb{G}) is a pair of correlated graphs sampled according to rule (2.1). Note that \mathcal{Q} is nothing but the marginal distribution on the first two coordinates of \mathcal{Q} . Our goal is to show $\text{TV}(\mathcal{P}, \mathcal{Q}) = o(1)$. To this end, we consider the likelihood ratio

$$L(G, \mathbb{G}) \triangleq \frac{d\mathcal{Q}}{d\mathcal{P}} \Big|_{(G, \mathbb{G})} = \frac{\sum_{\pi} \mathcal{Q}[G, \mathbb{G}, \pi]}{\mathcal{P}[G, \mathbb{G}]} = \frac{1}{n!} \sum_{\pi} \frac{\mathcal{Q}[G, \mathbb{G} \mid \pi]}{\mathcal{P}[G, \mathbb{G}]},$$

where $\mathcal{Q}[G, \mathbb{G} \mid \pi]$ is a short notation for $\mathcal{Q}[G, \mathbb{G} \mid \pi^* = \pi]$. Ideally we wish to show $\mathbb{E}_{\mathcal{Q}} L = \mathbb{E}_{\mathcal{P}} L^2 = 1 + o(1)$, but this fails due to the contribution from certain rare event. Therefore, we turn to the conditional moment for the likelihood ratio where we choose some “good” event \mathcal{G} measurable with respect to (G, \mathbb{G}, π^*) such that $\mathcal{Q}[\mathcal{G}] = 1 - o(1)$. Let $\mathcal{Q}'[\cdot] = \mathcal{Q}[\cdot \mid \mathcal{G}]$ and \mathcal{Q}' be the marginal distribution of the first two coordinates of \mathcal{Q}' .

We further assume that the marginal distribution of π^* under \mathcal{Q}' is still uniform, then the conditional likelihood ratio is then given by

$$L'(G, \mathbf{G}) \triangleq \frac{d\mathcal{Q}'}{d\mathbf{P}} \big|_{(G, \mathbf{G})} = \frac{\sum_{\pi} \mathcal{Q}'[G, \mathbf{G}, \pi]}{\mathbf{P}[G, \mathbf{G}]} = \frac{1}{n!} \sum_{\pi} \frac{\mathcal{Q}'[G, \mathbf{G} \mid \pi]}{\mathbf{P}[G, \mathbf{G}]}.$$

By the data processing inequality and the assumption that $\mathcal{Q}[\mathcal{G}] = 1 - o(1)$,

$$\text{TV}(\mathbf{Q}, \mathbf{Q}') \leq \text{TV}(\mathcal{Q}, \mathcal{Q}') = o(1).$$

As a result, once we show that $\mathbb{E}_{\mathbf{Q}'} L' = \mathbb{E}_{\mathbf{P}} L'^2 = 1 + o(1)$ for some appropriately chosen good event \mathcal{G} , then by the triangle inequality,

$$\text{TV}(\mathbf{P}, \mathbf{Q}) \leq \text{TV}(\mathbf{P}, \mathbf{Q}') + \text{TV}(\mathbf{Q}', \mathbf{Q}) \leq \sqrt{\log \mathbb{E}_{\mathbf{Q}'} L'} + o(1) = o(1),$$

where the second inequality follows from Pinsker's inequality and Jensen's inequality.

As we will see later, our good event \mathcal{G} will be measurable with respect to the isomorphic class of the π^* -intersection graph of G and \mathbf{G} , thus π^* does have uniform distribution under \mathcal{Q}' . In addition, in what follows, all the probability analysis conditioned on π^* is invariant with the realization of π^* . In particular, we can write

$$\mathbb{E}_{\mathbf{Q}'} L' = \mathbb{E}_{\pi^* \sim \mathcal{Q}'} \mathbb{E}_{(G, \mathbf{G}) \sim \mathcal{Q}'[\cdot \mid \pi^*]} L',$$

where $\mathbb{E}_{(G, \mathbf{G}) \sim \mathcal{Q}'[\cdot \mid \pi^*]} L'$ is invariant of the realization of π^* . As a result, for convenience of exposition, in what follows we regard π^* as certain fixed bijection from V to \mathbf{V} in order to avoid unnecessary complication from another layer of randomness.

We postpone the definition of our good event \mathcal{G} in Section 2.3.2, and we next investigate the conditional likelihood ration $L'(G, \mathbf{G})$ more carefully. For any bijection $\pi : V \rightarrow \mathbf{V}$, it is easy to see

$$\frac{\mathcal{Q}[G, \mathbf{G} \mid \pi]}{\mathbf{P}[G, \mathbf{G}]} = \prod_{(u, v) \in E_0} \ell(G_{u, v}, \mathbf{G}_{\pi(u), \pi(v)}), \quad (2.12)$$

where $\ell : \{0, 1\} \times \{0, 1\} \rightarrow \mathbb{R}$ denotes for the likelihood ratio function for a pair of edges

given by

$$\ell(x, y) = \begin{cases} \frac{1-2ps+ps^2}{(1-ps)^2} & x = y = 0; \\ \frac{1-s}{1-ps} & x = 1, y = 0 \text{ or } x = 0, y = 1; \\ \frac{1}{p} & x = y = 1. \end{cases} \quad (2.13)$$

Now suppose we are under H_1 . For a bijection $\pi : V \rightarrow \mathbf{V}$, define a permutation on V by $\sigma \triangleq \pi^{-1} \circ \pi^*$. Then π (respectively σ) induces a bijection Π from E_0 to \mathbf{E}_0 (respectively a permutation Σ on E_0), given by $\Pi((u, v)) = (\pi(u), \pi(v))$ (respectively $\Sigma((u, v)) = (\sigma(u), \sigma(v))$). For a given permutation σ on V , let \mathcal{O}_σ be the set of edge orbits induced by Σ , and we define

$$\mathcal{J}_\sigma = \{O \in \mathcal{O}_\sigma : G_{u,v} = \mathbf{G}_{\pi^*(u), \pi^*(v)} = 1, \text{ for all } (u, v) \in O\} \quad (2.14)$$

to be the set of edge orbits that are *entirely* contained in \mathcal{H}_{π^*} (recall (2.8) for the definition of π^* -intersection graph \mathcal{H}_{π^*}). Note that \mathcal{O}_σ is deterministic for a given σ , while \mathcal{J}_σ is random depending on π^* (which was assumed to be fixed) and the realization of (G, \mathbf{G}) . Let $H(\mathcal{J}_\sigma)$ be the subgraph of \mathcal{H}_{π^*} with vertices and edges from orbits in \mathcal{J}_σ and with slight abuse of notation, we denote by $|\mathcal{J}_\sigma|$ the number of edges in $H(\mathcal{J}_\sigma)$.

It is clear that once π^* and $\sigma = \pi^{-1} \circ \pi^*$ are fixed, the collections of $\{\ell(G_e, \mathbf{G}_{\Pi(e)}) : e \in O\}$ are mutually independent for $O \in \mathcal{O}_\sigma$. In addition, for $O \in \mathcal{J}_\sigma$ we have

$$\prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) = p^{-|O|}, \quad (2.15)$$

where $|O|$ denotes for the number of edges $e \in O$. The contribution to the untruncated moment from $O \in \mathcal{J}_\sigma$ blows up and thus calls for a truncation. Before doing that, we first prove the following lemma (similar to [45, Proposition 3]) which controls contribution from $O \notin \mathcal{J}_\sigma$.

Lemma 2.6. *Let $\pi = \pi^* \circ \sigma^{-1}$ (the notations for π^*, π, σ and Π, Σ are consistent as above). For $p, s \leq 0.1$ and $O \in \mathcal{O}_\sigma$,*

$$\mathbb{E}_{(G, \mathbf{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \left[\prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \mid O \notin \mathcal{J}_\sigma \right] \leq 1. \quad (2.16)$$

Proof. We first compute the expectation without conditioning on $O \notin \mathcal{J}_\sigma$ as follows:

$$\mathbb{E}_{(G, \mathbf{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) = \mathbb{E}_{(G, \mathbf{G}) \sim \mathbf{P}} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \ell(G_{\Sigma(e)}, \mathbf{G}_{\Pi(e)}).$$

The right hand side above can be further interpreted as $\text{Tr}(\mathcal{L}^{2|O|})$, where \mathcal{L} is the integral operator on the space of real functions on $\{0, 1\}$ induced by the kernel $\ell(\cdot, \cdot)$ in (2.13) as

$$(\mathcal{L}f)(x) \triangleq \mathbb{E}_{\mathbf{G}_e \sim \mathbf{P}}[\ell(x, \mathbf{G}_e)f(\mathbf{G}_e)] = \mathbb{E}_{(G_e, \mathbf{G}_e) \sim \mathbf{Q}}[f(\mathbf{G}_e) \mid G_e = x].$$

The matrix form of \mathcal{L} is given by $M(x, y) = \ell(x, y) \cdot \mathbf{P}[\mathbf{G}_e = y]$ for $(x, y) \in \{0, 1\} \times \{0, 1\}$, which can be written explicitly as

$$M = \begin{pmatrix} \frac{1-2ps+ps^2}{1-ps} & \frac{ps(1-s)}{1-ps} \\ 1-s & s \end{pmatrix}.$$

M has two eigenvalues 1 and $\rho \triangleq \frac{s(1-p)}{1-ps}$, so the unconditional expectation equals to $1 + \rho^{2|O|}$. See also [45, Proposition 1] for details.

Since conditioned on $O \notin \mathcal{J}_\sigma$ only excludes the case that $G_e = \mathbf{G}_{\Pi(e)} = 1$ for all $e \in O$,

$$\mathbb{E}_{(G, \mathbf{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \left[\prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \mid O \notin \mathcal{J}_\sigma \right] = \frac{1 + \rho^{2|O|} - s^{2|O|}}{1 - (ps^2)^{|O|}} \leq 1,$$

where the last inequality follows because for any $0 < p, s \leq 0.1$ and $k \geq 1$,

$$s^{2k} - \rho^{2k} = s^{2k} \left(1 - \left(\frac{1-p}{1-ps} \right)^{2k} \right) \geq s^{2k} (1 - (1-p)^k) \geq (ps^2)^k. \quad \square$$

We are now ready to derive the next lemma.

Lemma 2.7. *With notations in this subsection, we have*

$$\mathbb{E}_{(G, \mathbf{G}) \sim \mathbf{Q}'} L'(G, \mathbf{G}) \leq \frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{(G, \mathbf{G}, \pi^*) \sim \mathbf{Q}'} \frac{1}{n!} \sum_{\sigma} \frac{p^{-|\mathcal{J}_\sigma|}}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma]}. \quad (2.17)$$

Proof. First note that $\mathcal{Q}'[G, \mathbb{G}, \pi] \leq \frac{\mathcal{Q}[G, \mathbb{G}, \pi]}{\mathcal{Q}[\mathcal{G}]}$ holds for any triple (G, \mathbb{G}, π) . Thus,

$$\begin{aligned}
\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'} L'(G, \mathbb{G}) &= \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'} \sum_{\pi} \frac{\mathcal{Q}'[G, \mathbb{G}, \pi]}{\mathcal{P}[G, \mathbb{G}]} \\
&\leq \frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'} \sum_{\pi} \frac{\mathcal{Q}[G, \mathbb{G}, \pi]}{\mathcal{P}[G, \mathbb{G}]} \stackrel{(2.12)}{=} \frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'} \frac{1}{n!} \sum_{\pi} \prod_{e \in E_0} \ell(G_e, \mathbb{G}_{\Pi(e)}) \\
&= \frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{\pi^* \sim \mathcal{Q}'} \frac{1}{n!} \sum_{\sigma} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'[\cdot | \pi^*]} \prod_{O \in \mathcal{O}_{\sigma}} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}), \tag{2.18}
\end{aligned}$$

where in the last equity we changed from summation over π to summation over $\sigma = \pi^{-1} \circ \pi^*$. By Lemma 2.6, we can take conditional expectation with respect to \mathcal{J}_{σ} for each σ and obtain that

$$\begin{aligned}
&\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'[\cdot | \pi^*]} \prod_{O \in \mathcal{O}_{\sigma}} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \\
&\stackrel{(2.15)}{=} \mathbb{E}_{\mathcal{J}_{\sigma} \sim \mathcal{Q}'[\cdot | \pi^*]} \left[p^{-|\mathcal{J}_{\sigma}|} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'[\cdot | \pi^*, \mathcal{J}_{\sigma}]} \left[\prod_{O \in \mathcal{O}_{\sigma} \setminus \mathcal{J}_{\sigma}} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \mid \mathcal{J}_{\sigma} \right] \right]. \tag{2.19}
\end{aligned}$$

Note that for each fixed σ , and any realization J of \mathcal{J}_{σ} ,

$$\begin{aligned}
&\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}'[\cdot | \pi^*, \mathcal{J}_{\sigma} = J]} \left[\prod_{O \in \mathcal{O}_{\sigma} \setminus \mathcal{J}_{\sigma}} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \mid \mathcal{J}_{\sigma} = J \right] \\
&\leq \frac{1}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_{\sigma} = J]} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \left[\prod_{O \in \mathcal{O}_{\sigma} \setminus J} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \mid \mathcal{J}_{\sigma} = J \right], \tag{2.20}
\end{aligned}$$

where the term $\frac{1}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_{\sigma} = J]}$ emerges since we moved from \mathcal{Q}' to \mathcal{Q} . (Note that although \mathcal{Q} and \mathcal{Q}' are similar since presumably $\mathcal{Q}[\mathcal{G}] = 1 - o(1)$, the term $\frac{1}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_{\sigma} = J]}$ is not necessarily negligible since conditioned on $\mathcal{J}_{\sigma} = J$ may substantially decrease the probability for \mathcal{G} .)

By Lemma 2.6, we have that

$$\begin{aligned}
&\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \left[\prod_{O \in \mathcal{O}_{\sigma} \setminus J} \prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \mid \mathcal{J}_{\sigma} = J \right] \\
&= \prod_{O \in \mathcal{O}_{\sigma} \setminus J} \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot | \pi^*]} \left[\prod_{e \in O} \ell(G_e, \mathbb{G}_{\Pi(e)}) \mid O \notin \mathcal{J}_{\sigma} \right] \leq 1.
\end{aligned}$$

Combined with (2.18), (2.19) and (2.20), it yields that $\mathbb{E}_{(G, \mathbf{G}) \sim \mathbf{Q}'} L'(G, \mathbf{G})$ is upper-bounded by

$$\frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{\pi^* \sim \mathcal{Q}'} \frac{1}{n!} \sum_{\sigma} \mathbb{E}_{\mathcal{J}_{\sigma} \sim \mathcal{Q}'[\cdot | \pi^*]} \frac{p^{-|\mathcal{J}_{\sigma}|}}{\mathcal{Q}[\mathcal{G} | \pi^*, \mathcal{J}_{\sigma}]} = \frac{1}{\mathcal{Q}[\mathcal{G}]} \mathbb{E}_{(G, \mathbf{G}, \pi^*) \sim \mathcal{Q}'} \frac{1}{n!} \sum_{\sigma} \frac{p^{-|\mathcal{J}_{\sigma}|}}{\mathcal{Q}[\mathcal{G} | \pi^*, \mathcal{J}_{\sigma}]},$$

as required. \square

2.3.2 Truncation for the π^* -intersection graph

In light of Lemma 2.7, it suffice to show that the right hand side of (2.17) is upper-bounded by $1 + o(1)$ for some appropriately chosen event \mathcal{G} with $\mathcal{Q}[\mathcal{G}] = 1 - o(1)$. Under \mathbf{H}_1 , the π^* -intersection graph $\mathcal{H}_{\pi^*} = (V, \mathcal{E}_{\pi^*})$ of G and \mathbf{G} is an Erdős-Rényi graph $\mathbf{G}(n, \frac{\lambda}{n})$, where $\lambda = nps^2$. Recall that we are now under the assumption

$$p = n^{-\alpha+o(1)} \text{ for some } \alpha < 1 \text{ and } \lambda \leq \lambda_* - \varepsilon \text{ for some constant } \varepsilon > 0. \quad (2.21)$$

In this subsection, we will define our good event \mathcal{G} . To this end, we need some more notations. For simple graphs H and \mathcal{H} , a *labeled embedding* of H into \mathcal{H} is an injective map $\tau : H \rightarrow \mathcal{H}$, such that $(\tau(u), \tau(v))$ is an edge of \mathcal{H} when (u, v) is an edge of H . Further, we define an *unlabeled embedding* of H into \mathcal{H} to be an equivalent class of labeled embeddings: for two labeled embeddings $\tau_1, \tau_2 : H \rightarrow \mathcal{H}$, we say $\tau_1 \sim \tau_2$ if and only if there exists an automorphism $\phi : H \rightarrow H$, such that $\tau_2 = \tau_1 \circ \phi$.

Let $t(H, \mathcal{H})$ be the number of unlabeled embeddings of H into \mathcal{H} . Then it is clear that the number of labeled embeddings of H into \mathcal{H} is given by $\text{Aut}(H)t(H, \mathcal{H})$, where $\text{Aut}(H)$ stands for the number of automorphisms of H to itself. For each isomorphic class \mathcal{C} , pick a representative element $H_{\mathcal{C}} \in \mathcal{C}$ and fix it. For each $k \geq 2$, let $\mathfrak{C}_k = \mathfrak{C}_k(\mathcal{H})$ (respectively $\mathfrak{T}_k = \mathfrak{T}_k(\mathcal{H})$) be the collection of all such representatives $H_{\mathcal{C}}$ which are connected non-tree graphs (respectively trees) with k vertices so that $t(H_{\mathcal{C}}, \mathcal{H}) \geq 1$.

Denote $\xi = \frac{1}{2}[\varrho(\lambda_* - \varepsilon) + \varrho(\lambda_*)]$. Since $\lambda_* > 1$ in the case $\alpha < 1$, by Proposition 2.4 we have $\varrho(\lambda_* - \varepsilon) < \xi < \varrho(\lambda_*) = \frac{1}{\alpha}$. Hence, when n is large enough,

$$np^{\xi} \geq n^{\delta_0} \text{ for some constant } \delta_0 > 0. \quad (2.22)$$

Fix some positive constant $\delta < \delta_0$, we say that a graph $\mathcal{H} = (V, \mathcal{E})$ is *admissible* if it satisfies the following properties:

- (i) The maximal edge-vertex ratio over all subgraphs does not exceed ξ , i.e.

$$\max_{\emptyset \neq U \subset V} \frac{|\mathcal{E}(U)|}{|U|} \leq \xi. \quad (2.23)$$

- (ii) The maximal degree of \mathcal{H} is no more than $\log n$.
- (iii) Any connected subgraph containing at least two cycles has size larger than $2 \log \log n$.
- (iv) For any $k \geq 2$, the number of k -cycles is bounded by $n^{\delta k}$.

Define the good event $\mathcal{G} = \{\mathcal{H}_{\pi^*} \text{ is admissible}\}$.

Lemma 2.8. *For an Erdős-Rényi graph $\mathcal{H} \sim \mathbf{G}\left(n, \frac{\lambda}{n}\right)$,*

$$\mathbb{P}[\mathcal{H} \text{ is admissible}] = 1 - o(1). \quad (2.24)$$

In addition, there exists a constant $c = c(\lambda, \delta) > 0$, such that for any subgraph H satisfies $\mathbb{P}[\mathcal{H} \text{ is admissible} \mid H \subset \mathcal{H}] > 0$ and any event \mathcal{F} that is measurable with respect to and decreasing with edges that are not contained in H , we have

$$\mathbb{P}[\mathcal{H} \text{ is admissible} \mid H \subset \mathcal{H}, \mathcal{F}] \geq [1 - o(1)]c^{|E(H)|}. \quad (2.25)$$

Remark 2.9. By (2.24), we see that $\mathcal{Q}[\mathcal{G}] = 1 - o(1)$. In addition, conditioning on $\mathcal{J}_\sigma = J$ for some realization J sampled from $\mathcal{Q}[\cdot \mid \pi^*, \mathcal{G}]$, we have $\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma = J] > 0$. Since $\{\mathcal{J}_\sigma = J\} = \{H(J) \subset \mathcal{H}_{\pi^*}\} \cap \mathcal{F}$ where \mathcal{F} is the event that there is no other edge orbit (except those in J) that is entirely contained in \mathcal{H}_{π^*} , we see from (2.25) that

$$\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma = J] \geq [1 - o(1)]c^{|E(H(J))|} = [1 - o(1)]c^{|\mathcal{J}_\sigma|}. \quad (2.26)$$

The preceding inequality is useful for us since on the right hand side of (2.17) there is a term of $\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma]$.

Proof of Lemma 2.8. First we show (2.24). It suffices to bound the probability that either of (i)-(iv) fails. For (i), since $\varrho(\lambda) \leq \varrho(\lambda_* - \varepsilon) < \xi$, Proposition 2.2 gives $\mathbb{P}[(i) \text{ fails}] = o(1)$. $\mathbb{P}[(ii) \text{ or } (iii) \text{ fails}] = o(1)$ is well-known. Indeed, the typical value of the maximal degree in \mathcal{H} is of order $\frac{\log n}{\log \log n}$ (see e.g. [22, Theorem 3.4]), and the typical value for the minimal size of connected subgraphs containing at least two cycles in \mathcal{H} is at least of order $\log n$ (this can be shown by a simple union bound). For (iv), since the expected number of k -cycles in \mathcal{H} is bounded by λ^k , by Markov inequality we get that $\mathbb{P}[(iv) \text{ fails}] \leq \sum_{k=1}^{\infty} \frac{\lambda^k}{n^{\delta k}} = o(1)$. Altogether, this yields (2.24).

For (2.25), the case $H = \emptyset$ reduces to (2.24) by FKG inequality (since $\{\mathcal{H} \text{ is admissible}\}$ is a decreasing event), so we can assume $H \neq \emptyset$. The condition $\mathbb{P}[\mathcal{H} \text{ is admissible} \mid H \subset \mathcal{H}] > 0$ implies that the subgraph H satisfies all conditions in admissibility. Let V_1 be the vertex set of H , and $V_2 = V \setminus V_1$. Consider the following three events:

$$\begin{aligned} \mathcal{A}_1 &= \{\text{There is no edge within } V_1 \text{ except those in } H, \mathcal{E}(V_1, V_2) = \emptyset\}, \\ \mathcal{A}_2 &= \{\text{There is no cycle in } V_2 \text{ with length less than } K \triangleq \lceil \delta^{-1} \rceil\}, \\ \mathcal{A}_3 &= \{\text{The subgraph in } V_2 \text{ satisfies (i), (ii), (iii) in admissibility}\}. \end{aligned}$$

We claim that conditioned on $H \subset \mathcal{H}$, we have \mathcal{H} is admissible as long as $\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3$ holds (we comment that the purpose of defining event \mathcal{A}_2 is to handle the potential scenario where there are $n^{\delta k}$ k -cycles in H). Clearly, \mathcal{H} satisfies (i), (ii), (iii) in admissibility by $\mathcal{A}_1 \cap \mathcal{A}_3$. For (iv), the case $k < K$ is guaranteed by \mathcal{A}_2 . When $K \leq k \leq \log \log n$, no two k -cycle share a common vertex by (iii), thus the number of k -cycles is at most $n/k < n^{\delta k}$ since $k \geq K$; and when $k \geq \log \log n$, the number of k -cycles is bounded by $n(\log n)^k \leq n^{\delta k}$ from (ii). Thus (iv) also holds and \mathcal{H} is admissible.

Note that \mathcal{A}_1 is independent with $\mathcal{A}_2 \cap \mathcal{A}_3$, $\mathbb{P}[\mathcal{A}_1] \geq e^{-2\lambda|E(H)|}$, and $\mathbb{P}[\mathcal{A}_2] > C_\delta$ for some constant $C_\delta > 0$ since the distribution of small cycles are approximately independent Poisson variables (one can also use FKG inequality instead of approximate independence here since the number of small cycles are all increasing with the graph). Since the subgraph within V_2 is an Erdős-Rényi, we get $\mathbb{P}[\mathcal{A}_3^c] = o(1)$ from (2.24). Therefore, for some small

constant $c = c(\lambda, \delta) > 0$ we have

$$\begin{aligned} \mathbb{P}[\mathcal{H} \text{ is admissible} \mid H \subset \mathcal{H}, \mathcal{F}] &\geq \mathbb{P}[\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3 \mid \mathcal{F}] \geq \mathbb{P}[\mathcal{A}_1 \cap \mathcal{A}_2 \cap \mathcal{A}_3] \\ &\geq \mathbb{P}[\mathcal{A}_1] (\mathbb{P}[\mathcal{A}_2] - \mathbb{P}[\mathcal{A}_3^c]) \geq [C_\delta - o(1)] e^{-2\lambda|E(H)|} \geq [1 - o(1)] c^{|E(H)|}, \end{aligned}$$

where we applied FKG inequality for the last transition in the first line (note that $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ are all decreasing events). This completes the proof of the lemma. \square

As a result of admissibility, we have the following bounds on subgraph counts.

Lemma 2.10. *For an admissible graph \mathcal{H} and for $k \geq 2$, we have that the total number of labeled embeddings of $T \in \mathfrak{T}_k$ is bounded by $n(4 \log n)^{2(k-1)}$, i.e.*

$$\sum_{T \in \mathfrak{T}_k} \text{Aut}(T) t(T, \mathcal{H}) \leq n(4 \log n)^{2(k-1)}. \quad (2.27)$$

In addition, the total number of labeled embeddings of $C \in \mathfrak{C}_k$ is bounded by $k^3(2^{\xi+1}n^\delta)^k$, i.e.

$$\sum_{C \in \mathfrak{C}_k} \text{Aut}(C) t(C, \mathcal{H}) \leq k^3(2^{\xi+1}n^\delta)^k. \quad (2.28)$$

Proof. By [33], the number for isomorphic classes of trees with k vertices is at most 4^{k-1} . For each class, we claim that the number of labeled embedding is at most $n(\log n)^{2(k-1)}$. This is because each embedding can be encoded by a walk path of length $2(k-1)$ on the graph which corresponds to the depth-first search contour of the image of the embedding. By (ii) in admissibility, the number of paths of length $2(k-1)$ is at most $n(\log n)^{2(k-1)}$. This yields (2.27).

For (2.28), note that every labeled non-tree subgraph with k vertices on \mathcal{H} can be constructed by the following steps:

Step 1. Pick an isomorphic class of connected graphs with k vertices and k edges, and take its representative C with vertices labeled by v_1, \dots, v_k .

Step 2. Choose a labeled embedding $\tau : C \rightarrow \mathcal{H}$.

Step 3. Add some of the remaining edges within $\{\tau(v_1), \dots, \tau(v_k)\}$ to $\tau(C)$ and get the final subgraph.

The number of isomorphic classes in **Step 1** is no more than $k^2 4^{k-1}$ since we can first pick a tree of k vertices and then add an extra edge. For any connected subgraph C with k vertices and k edges, C is a union of a cycle with $r \leq k$ vertices together with some trees. The number of labeled embeddings of the cycle is bounded by $2r(n^\delta)^r$ by (iv) in admissibility, and once this is done, the number of ways to embed the rest of trees is bounded by $(\log n)^{2(k-r)} \leq n^{\delta(k-r)}$ from (ii) in admissibility (and a similar argument as for (2.27)). So, the number of labeled embedding in **Step 2** is bounded by $2k(n^\delta)^k$. Finally, for any labeled embedding $\tau : C \rightarrow \mathcal{H}$, since (i) holds, the total number of remaining edges between $\tau(v_1), \dots, \tau(v_k)$ is bounded by $(\xi - 1)k$, we see that the number of choices for **Step 3** is at most $2^{(\xi-1)k}$. Now a simple application of multiplication rule yields (2.28). \square

2.3.3 The truncated exponential moment

We now prove the following bound.

Proposition 2.11. *Suppose (2.22) holds and suppose that \mathcal{H} is admissible. Then as $n \rightarrow \infty$,*

$$\frac{1}{n!} \sum_{\sigma} \frac{p^{-|\mathcal{J}_{\sigma}|}}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_{\sigma}]} \leq 1 + o(1). \quad (2.29)$$

Combined with Lemmas 2.7 and 2.8 as well as discussions at the beginning of Section 2.3.1, this then completes the proof of (2.3).

When proving (2.29), it would be convenient to first fix a subgraph H and sum over all permutations σ with $H(\mathcal{J}_{\sigma}) = H$, and then sum over all possible H . To this end, we will need the following lemma on an upper bound for the number of permutations σ on V such that $H(\mathcal{J}_{\sigma}) \cong H_C$ for each possible isomorphic class representative H_C that can arise from an admissible graph and some permutation. Note that for each realization of $H(\mathcal{J}_{\sigma})$, the collection of its components is isomorphic to a union of some C 's in $\bigcup \mathfrak{C}_r$ and some T 's in $\bigcup \mathfrak{T}_s$, since $H(\mathcal{J}_{\sigma})$ is always a subgraph of \mathcal{H} .

Lemma 2.12. *Suppose the collection of components of H_C is isomorphic to a union of some $C_i \in \mathfrak{C}_{r_i}$ with x_i copies for $1 \leq i \leq l$ and some $T_j \in \mathfrak{T}_{s_j}$ with y_j copies for $1 \leq j \leq m$, where $C_1, \dots, C_l, T_1, \dots, T_m$ are in distinct isomorphic classes. Then*

$$|\sigma : H(\mathcal{J}_\sigma) \cong H_C| \leq (n - \sum_{i \in [l]} x_i r_i - \sum_{j \in [m]} y_j s_j)! \times \prod_{i \in [l]} (\text{Aut}(C_i) t(C_i, \mathcal{H}))^{x_i} \prod_{j \in [m]} (\text{Aut}(T_j) t(T_j, \mathcal{H}))^{y_j}. \quad (2.30)$$

Proof. First we choose an unlabeled embedding τ for H_C in \mathcal{H} . Since $C_1, \dots, C_l, T_1, \dots, T_m$ are in distinct isomorphic classes, τ can be viewed as a product of $\tau|_{x_i \cdot C_i}$ and $\tau|_{y_j \cdot T_j}$ for $1 \leq i \leq l$ and $1 \leq j \leq m$, where each $\tau|_{x_i \cdot C_i}$ (respectively $\tau|_{y_j \cdot T_j}$) is an unlabeled embedding for x_i disjoint copies of C_i (respectively y_j disjoint copies of T_j). Therefore, the number of choices for τ is bounded by

$$\prod_{i \in [l]} \binom{t(C_i, \mathcal{H})}{x_i} \prod_{j \in [m]} \binom{t(T_j, \mathcal{H})}{y_j} \leq \prod_{i \in [l]} \frac{t(C_i, \mathcal{H})^{x_i}}{x_i!} \prod_{j \in [m]} \frac{t(T_j, \mathcal{H})^{y_j}}{y_j!}. \quad (2.31)$$

For each unlabeled embedding τ , we wish to bound the number of σ such that $H(\mathcal{J}_\sigma) = \phi \circ \tau(H_C)$ for some permutation ϕ on the vertex set of $\tau(H_C)$, where the $=$ is in the sense of equal for labeled graphs. We claim that the number of such σ 's is bounded by

$$\left(n - \sum_{i \in [l]} x_i r_i - \sum_{j \in [m]} y_j s_j \right)! \times \prod_{i \in [l]} x_i! \text{Aut}(C_i)^{x_i} \prod_{j \in [m]} y_j! \text{Aut}(T_j)^{y_j}. \quad (2.32)$$

Let V_1 be the vertex set of $\tau(H_C)$, then $|V_1| = \sum_{i \in [l]} x_i r_i + \sum_{j \in [m]} y_j s_j$. It is clear that any aforementioned desired σ can be decomposed into two permutations σ_1 and σ_2 on V_1 and $V \setminus V_1$, respectively. The number of choices for σ_2 is at most $(n - |V_1|)!$ (it may be strictly less than $(n - |V_1|)!$ since on $V \setminus V_1$ we are not allowed to produce another edge orbit that is entirely in \mathcal{H}). In order to bound the number of choices for σ_1 , we use the following crucial observation: for any $u \in V_1$, we have σ_1 inhibits to an isomorphism between the two components of $\tau(H_C)$ which contain u and $\sigma_1(u)$. That is to say, for any u adjacent to v in $\tau(H_C)$, $\sigma(u)$ is also adjacent to $\sigma(u)$ in $\tau(H_C)$; similarly, for any w

adjacent to z in $\tau(H_C)$, $\sigma^{-1}(w)$ is also adjacent to $\sigma^{-1}(z)$ in $\tau(H_C)$. This is true because of the definition of edge orbit and our requirement that $H(\mathcal{J}_\sigma)$ is entirely contained in \mathcal{H} . From this observation, for each C_i (and similarly for T_j) we will “permute” its x_i copies so that σ maps one copy to its image under the permutation, and within each copy of C_i we have the freedom of choosing an arbitrary automorphism. In addition, the choice of such permutations and automorphisms completely determine σ_1 . Therefore, the number of choices for σ_1 is bounded by $\prod_{i \in [l]} x_i! \text{Aut}(C_i)^{x_i} \prod_{j \in [m]} y_j! \text{Aut}(T_j)^{y_j}$. This proves (2.32). Combined with (2.31), it yields (2.30). \square

Proof of Proposition 2.11. We have

$$\frac{1}{n!} \sum_{\sigma} \frac{p^{-|\mathcal{J}_\sigma|}}{\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma]} \leq \frac{1}{n!} \sum \frac{p^{-|E(H_C)|}}{\mathcal{Q}[\mathcal{G} \mid \pi^*, H(\mathcal{J}_\sigma) \cong H_C]} \times |\sigma : H(\mathcal{J}_\sigma) \cong H_C|, \quad (2.33)$$

where the sum is over all possible representatives H_C . We will bound $\mathcal{Q}[\mathcal{G} \mid \pi^*, \mathcal{J}_\sigma]$ by (2.26). For each tree component with s vertices it is clear that the number of edges is $s - 1$, and crucially for each non-tree component with r vertices, we use (2.23) to bound the number of edges in the component. In addition, we use (2.30) to bound $|\sigma : H(\mathcal{J}_\sigma) \cong H_C|$. Therefore, by enumerating all the possible isomorphic class representatives H_C , we can upper-bound (2.33) by

$$\begin{aligned} & \frac{1}{n!} \sum_{l, m \geq 0} \sum_{C_1, \dots, C_l \in \bigcup \mathfrak{C}_r} \sum_{T_1, \dots, T_m \in \bigcup \mathfrak{T}_s} \sum_{x_1, \dots, x_l > 0} \sum_{y_1, \dots, y_m > 0} [1 + o(1)] \\ & \times (cp)^{-\xi \sum_{i \in [l]} x_i |C_i| - \sum_{j \in [m]} y_j (|T_j| - 1)} \left(n - \sum_{i \in [l]} x_i |C_i| - \sum_{j \in [m]} y_j |T_j| \right)! \\ & \times \prod_{i \in [l]} (\text{Aut}(C_i) t(C_i, \mathcal{H}))^{x_i} \prod_{j \in [m]} (\text{Aut}(T_j) t(T_j, \mathcal{H}))^{y_j}, \end{aligned}$$

where $c = c(\lambda, \delta)$ is the constant in (2.26). By Stirling's formula we see

$$\frac{(n - k)!}{n!} \leq \left(\frac{2e}{n} \right)^k, \text{ for all } 0 \leq k \leq n.$$

Write $c' = c/2e$, then (3.68) can be further bounded by $1 + o(1)$ multiples

$$\begin{aligned}
& \sum_{l \geq 0} \sum_{C_1, \dots, C_l \in \bigcup \mathfrak{C}_r} \sum_{x_1, \dots, x_l > 0} \left(\frac{1}{c' n p^\xi} \right)^{x_1 |C_1| + \dots + x_l |C_l|} \prod_{i \in [l]} (\text{Aut}(C_i) t(C_i, \mathcal{H}))^{x_i} \\
& \times \sum_{m \geq 0} \sum_{T_1, \dots, T_m \in \bigcup \mathfrak{T}_s} \sum_{y_1, \dots, y_m > 0} \left(\frac{1}{c' n p} \right)^{y_1 |T_1| + \dots + y_m |T_m|} \prod_{j \in [m]} (p \text{Aut}(T_j) t(T_j, \mathcal{H}))^{y_j} \\
& \leq \prod_{C \in \bigcup \mathfrak{C}_r} \left(1 + \sum_{x > 0} \left(\frac{\text{Aut}(C) t(C, \mathcal{H})}{(c' n p^\xi)^{|C|}} \right)^x \right) \prod_{T \in \bigcup \mathfrak{T}_s} \left(1 + \sum_{y > 0} \left(\frac{p \text{Aut}(T) t(T, \mathcal{H})}{(c' n p)^{|T|}} \right)^y \right). \quad (2.34)
\end{aligned}$$

Under the assumption that \mathcal{H} is admissible and the condition $np \geq np^\xi \geq n^{\delta_0}$ by (2.22), we get from (2.27) that

$$\frac{p \text{Aut}(T) t(T, \mathcal{H})}{(c' n p)^{|T|}} \leq \frac{np(4 \log n)^{2(|T|-1)}}{c' n p \cdot (c' n^{\delta_0})^{|T|-1}} = o(1), \quad \text{for all } T \in \mathfrak{T}_s. \quad (2.35)$$

Similarly, we get from (2.28) that (recall $\delta < \delta_0$ by choice)

$$\frac{\text{Aut}(C) t(C, \mathcal{H})}{(c' n p^\xi)^{|C|}} \leq \frac{|C|^3 (2^{\xi+1} n^\delta)^{|C|}}{(c' n^{\delta_0})^{|C|}} = o(1), \quad \text{for all } C \in \mathfrak{C}_r. \quad (2.36)$$

Since $\log(1+x) \leq x$ for all $x > 0$, we get from (2.35) and (2.36) that the logarithm of (3.69) is bounded by $o(1)$ plus

$$[1 + o(1)] \left[\sum_{C \in \bigcup \mathfrak{C}_r} \frac{\text{Aut}(C) t(C, \mathcal{H})}{(c' n p^\xi)^{|C|}} + \sum_{T \in \bigcup \mathfrak{T}_s} \frac{p \text{Aut}(T) t(T, \mathcal{H})}{(c' n p)^{|T|}} \right]. \quad (2.37)$$

In order to bound (2.37), we first sum over all $C \in \mathfrak{C}_r$ and $T \in \mathfrak{T}_s$ then sum over r, s .

Applying this procedure and using (2.27) and (2.28), we get that (2.37) is at most

$$[1 + o(1)] \left[\sum_{k \geq 2} \frac{k^3 (2^{\xi+1} n^\delta)^k}{(c' n^{\delta_0})^k} + \sum_{k \geq 2} \frac{np(4 \log n)^{k-1}}{(c' n p)^k} \right] = o(1),$$

which shows that the logarithm of (3.69) is $o(1)$. This implies that (3.68) is bounded by $1 + o(1)$. Combined with (2.33), this completes the proof of Proposition 2.11. \square

References

- [1] *HLT '05: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, USA, 2005. Association for Computational Linguistics.
- [2] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.
- [3] V. Anantharam and J. Salez. The densest subgraph problem in sparse random graphs. *Ann. Appl. Probab.*, 26(1):305–327, 2016.
- [4] B. Barak, C.-N. Chou, Z. Lei, T. Schramm, and Y. Sheng. (nearly) efficient algorithms for the graph matching problem on correlated random graphs. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [5] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 26–33 vol. 1, 2005.
- [6] M. Bozorg, S. Salehkaleybar, and M. Hashemi. Seedless graph matching via tail of degree distribution for correlated Erdős-Rényi graphs. Preprint, arXiv:1907.06334.
- [7] J. A. Cain, P. Sanders, and N. Wormald. The random graph threshold for k -orientability and a fast algorithm for optimal multiple-choice allocation. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 469–476. ACM, New York, 2007.
- [8] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006.

- [9] D. Cullina and N. Kiyavash. Exact alignment recovery for correlated Erdős-Rényi graphs. Preprint, arXiv:1711.06783.
- [10] D. Cullina and N. Kiyavash. Improved achievability and converse bounds for erdos-renyi graph matching. In *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, SIGMETRICS '16, page 6372, New York, NY, USA, 2016. Association for Computing Machinery.
- [11] D. Cullina, N. Kiyavash, P. Mittal, and H. V. Poor. Partial recovery of Erdős-Rényi graph alignment via k -core alignment. SIGMETRICS '20, page 99100, New York, NY, USA, 2020. Association for Computing Machinery.
- [12] O. E. Dai, D. Cullina, N. Kiyavash, and M. Grossglauser. Analysis of a canonical labeling algorithm for the alignment of correlated Erdős-Rényi graphs. *Proc. ACM Meas. Anal. Comput. Syst.*, 3(2), jun 2019.
- [13] J. Ding, Z. Ma, Y. Wu, and J. Xu. Efficient random graph matching via degree profiles. *Probab. Theory Related Fields*, 179(1-2):29–115, 2021.
- [14] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.
- [15] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations II: Erdős-Rényi graphs and universality. Preprint, arXiv:1907.08883.
- [16] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations: Algorithm and theory. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2985–2995. PMLR, 13–18 Jul 2020.
- [17] S. Feizi, G. Quon, M. Medard, M. Kellis, and A. Jadbabaie. Spectral alignment of networks. Preprint, arXiv:1602.04181.

- [18] D. Fernholz and V. Ramachandran. The k -orientability thresholds for $G_{n,p}$. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 459–468. ACM, New York, 2007.
- [19] N. Fountoulakis, M. Khosla, and K. Panagiotou. The multiple-orientability thresholds for random hypergraphs. *Combin. Probab. Comput.*, 25(6):870–908, 2016.
- [20] A. Frieze and M. Karoński. *Introduction to random graphs*. available at <https://www.math.cmu.edu/~af1p/BOOK.pdf>.
- [21] L. Ganassali and L. Massoulié. From tree matching to sparse graph alignment. In J. Abernethy and S. Agarwal, editors, *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1633–1665. PMLR, 09–12 Jul 2020.
- [22] P. Gao and N. C. Wormald. Load balancing and orientability thresholds for random hypergraphs [extended abstract]. In *STOC’10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*, pages 97–103. ACM, New York, 2010.
- [23] B. Hajek. Performance of global load balancing by local adjustment. *IEEE Trans. Inform. Theory*, 36(6):1398–1414, 1990.
- [24] G. Hall and L. Massoulié. Partial recovery in the graph alignment problem. Preprint, arXiv:2007.00533.
- [25] E. Kazemi, S. H. Hassani, and M. Grossglauser. Growing a graph matching from a handful of seeds. *Proc. VLDB Endow.*, 8(10):10101021, jun 2015.
- [26] V. Lyzinski, D. E. Fishkind, and C. E. Priebe. Seeded graph matching for correlated Erdős-Rényi graphs. *J. Mach. Learn. Res.*, 15:3513–3540, 2014.

- [27] C. Mao, M. Rudelson, and K. Tikhomirov. Exact matching of random graphs with constant correlation. Preprint, arXiv:2110.05000.
- [28] C. Mao, Y. Wu, J. Xu, and S. H. Yu. Testing network correlation efficiently via counting trees. Preprint, arXiv:2110.11816.
- [29] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, USA, 2005.
- [30] E. Mossel and J. Xu. Seeded graph matching via large neighborhood statistics. *Random Structures Algorithms*, 57(3):570–611, 2020.
- [31] A. Narayanan and V. Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, 2008.
- [32] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187, 2009.
- [33] R. Otter. The number of trees. *Ann. of Math. (2)*, 49:583–599, 1948.
- [34] P. Pedarsani and M. Grossglauser. On the privacy of anonymized networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, page 12351243, New York, NY, USA, 2011. Association for Computing Machinery.
- [35] B. Pittel. On tree census and the giant component in sparse random graphs. *Random Structures Algorithms*, 1(3):311–342, 1990.
- [36] M. Z. Racz and A. Sridhar. Correlated randomly growing graphs. to appear in *Ann. Appl. Probab.*
- [37] M. Z. Racz and A. Sridhar. Correlated stochastic block models: Exact graph matching with applications to recovering communities. In *Advances in Neural Information Processing Systems*, 2021.

- [38] F. Shirani, S. Garg, and E. Erkip. Seeded graph matching: Efficient algorithms and theoretical guarantees. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 253–257, 2017.
- [39] R. Singh, J. Xu, and B. Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences of the United States of America*, 105:12763–8, 10 2008.
- [40] J. T. Vogelstein, J. M. Conroy, V. Lyzinski, L. J. Podrazik, S. G. Kratzer, E. T. Harley, D. E. Fishkind, R. J. Vogelstein, and C. E. Priebe. Fast approximate quadratic programming for graph matching. *PLOS ONE*, 10(4):1–17, 04 2015.
- [41] H. Wang, Y. Wu, J. Xu, and I. Yolou. Random graph matching in geometric models: the case of complete graphs. Preprint, arXiv:2202.10662.
- [42] Y. Wu, J. Xu, and S. H. Yu. Settling the sharp reconstruction thresholds of random graph matching. Preprint, arXiv:2102.00082.
- [43] Y. Wu, J. Xu, and S. H. Yu. Testing correlation of unlabeled random graphs. Preprint, arXiv:2008.10097.
- [44] L. Yartseva and M. Grossglauser. On the performance of percolation graph matching. In *Proceedings of the First ACM Conference on Online Social Networks, COSN '13*, page 119130, New York, NY, USA, 2013. Association for Computing Machinery.

3 Informational phase transition of partial recovery in the correlated random graph model

For two correlated graphs which are independently sub-sampled from a common Erdős-Rényi graph $\mathbf{G}(n, p)$, we wish to recover their *latent* vertex matching from the observation of these two graphs *without labels*. When $p = n^{-\alpha+o(1)}$ for $\alpha \in (0, 1]$, we establish a sharp information-theoretic threshold for whether it is possible to correctly match a positive fraction of vertices. Our result sharpens a constant factor in a recent work by Wu, Xu and Yu. This section is based on a joint work with Jian Ding.

3.1 Introduction

In this paper, we study the information-theoretic threshold for recovering the latent matching between two correlated Erdős-Rényi graphs. To mathematically make sense of the problem, we first need to choose a model for a pair of correlated Erdős-Rényi graphs, and a natural choice is that the two graphs are independently sub-sampled from a common Erdős-Rényi graph. More precisely, for two vertex sets V and \mathbf{V} with cardinality n , let E_0 be the set of unordered pairs (u, v) with $u, v \in V, u \neq v$ and define \mathbf{E}_0 similarly with respect to \mathbf{V} . For some model parameters $p, s \in (0, 1)$, define \mathbf{Q} to be the law for two (correlated) random graphs $G = (V, E)$ and $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ generated as follow: sample a uniform bijection $\pi^* : V \rightarrow \mathbf{V}$, independent Bernoulli variables $I_{(u,v)}$ with parameter p for $(u, v) \in E_0$ as well as independent Bernoulli variables $J_{(u,v)}, \mathbf{J}_{(\mathbf{u}, \mathbf{v})}$ with parameter s for $(u, v) \in E_0$ and $(\mathbf{u}, \mathbf{v}) \in \mathbf{E}_0$. Let

$$G_{(u,v)} = I_{(u,v)} J_{(u,v)}, \forall (u, v) \in E_0, \quad \mathbf{G}_{(\mathbf{u}, \mathbf{v})} = I_{((\pi^*)^{-1}(\mathbf{u}), (\pi^*)^{-1}(\mathbf{v}))} \mathbf{J}_{(\mathbf{u}, \mathbf{v})}, \forall (\mathbf{u}, \mathbf{v}) \in \mathbf{E}_0, \quad (3.1)$$

and $E = \{e \in E_0 : G_e = 1\}, \mathbf{E} = \{\mathbf{e} \in \mathbf{E}_0 : \mathbf{G}_{\mathbf{e}} = 1\}$. It is obvious that marginally G is an Erdős-Rényi graph with edge density p (which we denote as $\mathbf{G}(n, p)$) and so is \mathbf{G} .

A fundamental question is to recover the *latent* matching π^* from the observation of (G, \mathbf{G}) . More precisely, we wish to find an estimator $\hat{\pi}$ which is measurable with respect

to (G, \mathbf{G}) that maximizes $\text{overlap}(\pi^*, \hat{\pi})$, where $\text{overlap}(\pi^*, \hat{\pi}) = |\{v \in V : \pi^*(v) = \hat{\pi}(v)\}|$. Our main contribution is a sharp information-theoretic threshold for partial recovery, i.e., whether there exists a $\hat{\pi}$ such that $\text{overlap}(\pi^*, \hat{\pi}) \geq \delta n$ for some positive constant δ .

Theorem 3.1. *Suppose $p = n^{-\alpha+o(1)}$ for some fixed $\alpha \in (0, 1]$ (where $o(1)$ denotes a term vanishing in n). Let $\lambda_* = \varrho^{-1}(\frac{1}{\alpha})$ (here ϱ^{-1} is defined in Proposition 3.2 below) and $\lambda = np s^2$. Then for any positive constant ε the following hold:*

- *If $\lambda \geq \lambda_* + \varepsilon$, then there exist an estimator $\hat{\pi}$ and a positive constant $\delta = \delta(\alpha, \varepsilon)$, such that*

$$\mathbb{P}[\text{overlap}(\pi^*, \hat{\pi}) \geq \delta n] = 1 - o(1). \quad (3.2)$$

- *If $\alpha < 1$ and $\lambda \leq \lambda_* - \varepsilon$, then for any estimator $\hat{\pi}$ and any positive constant δ ,*

$$\mathbb{P}[\text{overlap}(\pi^*, \hat{\pi}) \geq \delta n] = o(1). \quad (3.3)$$

We emphasize that (3.3) in the case of $\alpha = 1$ was shown in [44]. While our proof should also be able to cover this case, we choose to focus on the case of $\alpha < 1$ as this assumption helps avoiding some technical complications. As an important contribution, [44] established a sharp threshold for $\alpha = 0$ and up-to-constant upper and lower bounds for $0 < \alpha \leq 1$. So in summary, our work is hugely inspired by [44] and in return sharpens a constant factor therein and thus fills the remaining gap in the regime for $0 < \alpha \leq 1$. In addition, we note that the sharp threshold for exact recovery was established in [44] which concerns the existence of $\hat{\pi}$ such that $\hat{\pi} = \pi^*$.

3.1.1 Background and related results

Recently, there has been extensive study on the problem of matching the vertex correspondence between two correlated graphs and the closely related problem of detecting the correlation between two graphs. On the one hand, questions of this type have been raised from various applied fields such as social network analysis [30, 31], computer vision

[9, 3], computational biology [37, 38] and natural language processing [20]; on the other hand, graph matching problems seem to provide another important set of examples which exhibit the intriguing *information-computation gap*, whose theoretical analysis integrates tools from various branches of mathematics.

Despite the fact that Erdős-Rényi Graph perhaps does not quite capture important features for any network arising from realistic problems, it is nevertheless reasonable to start our (presumably long) journey of completely understanding the information-computation phase transition for graph matching problems from a clean, simple and in some sense canonical random graph model such as Erdős-Rényi. Along this line, many progress has been made recently, including information-theoretic analysis [8, 7, 19, 45, 44] and proposals for various efficient algorithms [36, 46, 25, 23, 17, 38, 3, 13, 5, 9, 10, 32, 16, 20, 15, 27, 28]. As of now, it seems fair to say that we are still relatively far away from being able to completely understand the phase transition for computational complexity of graph matching problems. As in many other problems of this type, an information-theoretic phase transition is easier and usually will also guide the study on the transition for computational complexity. Together with previous works [45, 44, 12] (which were naturally inspired by earlier works such as [8, 7, 19]), it seems now we have achieved a fairly satisfying understanding on the information-theoretic transition and we hope that this may be of help for future study on computational aspects.

As hinted from earlier discussions, an important (and in fact a substantially more important) research direction is to study graph matching problems on more realistic graph models other than Erdős-Rényi. This ambitious program has started seeing some progress, sometimes paralleling to that on Erdős-Rényi graphs and sometimes inspired by insights accumulated on Erdős-Rényi. For instance, a model for correlated randomly growing graphs was studied in [35], graph matching for correlated stochastic block model was studied in [36] and graph matching for correlated random geometric graphs was studied in [40]. In a very recent work [9], a related matching problem (albeit somewhat different from graph

matching) was studied and it seems the method developed therein enjoyed direct and successful applications to single-cell problems.

3.1.2 Connection to previous works

We have learned from [45, 44] important insights on information thresholds for graph matching problems. An additional ingredient we realized is the connection to the densest subgraph, which allowed us to improve [45] and establish the sharp detection threshold as in [12]. The densest subgraph problem arose in the study of load balancing problem [25] and much progress has been made on densest subgraphs for random graphs [8, 20, 24, 21]. Of particular importance to us is the work of [3] which in particular established the asymptotic value for the maximal subgraph density of an Erdős-Rényi graph using the objective method from [2], as incorporated in the next proposition.

Proposition 3.2. (*[3, Theorems 1 and 3], see also [12, Propositions 2.1 and 2.3]*) *There exists a continuous, strictly increasing and unbounded function $\varrho : [1, \infty) \rightarrow [1, \infty)$ (which can be explicitly characterized via a variational problem) with $\varrho(1) = 1$, such that for any $\lambda \geq 1$ the maximal edge-vertex ratio over all nonempty subgraphs of an Erdős-Rényi graph $\mathcal{H} = (V, \mathcal{E}) \sim \mathbf{G}(n, \frac{\lambda}{n})$ concentrates around $\varrho(\lambda)$ as $n \rightarrow \infty$, i.e.,*

$$\max_{\emptyset \neq U \subset V} \frac{|\mathcal{E}(U)|}{|U|} \rightarrow \varrho(\lambda) \text{ in probability as } n \rightarrow \infty. \quad (3.4)$$

Furthermore, when $\lambda > 1$, there is a constant $c_\lambda > 0$ such that with probability tending to 1 as $n \rightarrow \infty$, the densest subgraph in \mathcal{H} (i.e. the maximizer of the left hand side of (3.4)) has size at least $c_\lambda n$.

Denote $\varrho^{-1} : [1, \infty) \rightarrow [1, \infty)$ for the inverse function of ϱ . Building on insights from [45] and using Proposition 3.2, we have established the following detection threshold in [12] as stated in the next proposition. Define \mathbf{P} as the law of a pair of independent Erdős-Rényi graphs $\mathbf{G}(n, ps)$ on V and \mathbf{V} , respectively. For two probability measures μ and ν , we denote by $\text{TV}(\mu, \nu) = \sup_A (\mu(A) - \nu(A))$ the total variation distance between μ and ν .

Proposition 3.3. ([\[12, Theorem 1.1\]](#)) *With notations in Theorem [3.1](#), the following hold:*

- (i) *If $\lambda \geq \lambda_* + \varepsilon$, then $\text{TV}(\mathbf{P}, \mathbf{Q}) = 1 - o(1)$;*
- (ii) *If $\alpha < 1$ and $\lambda \leq \lambda_* - \varepsilon$, then $\text{TV}(\mathbf{P}, \mathbf{Q}) = o(1)$.*

In order to prove Theorem [3.1](#), we combine insights from [\[44\]](#) and the proof of Proposition [3.3](#). It is perhaps not surprising that the partial recovery threshold coincides with the detection threshold. On the contrary, what may be unexpected at the first glance is that given [\[12\]](#) it still requires substantial amount of non-trivial work to prove impossibility of partial recovery as the default folklore assumption is that detection is easier than recovery. Indeed, detection is easier than exact recovery since with possibility of exact recovery one should be able to detect the correlation by examining the intersection of the two graphs under this (estimated) matching. But if the estimator only achieves partial recovery, the intersection graph is not necessarily “bigger” than that of a random matching, which partly explains the difficulty for proving impossibility of partial recovery. As an analogy, the difficulty we face is similar to the challenge addressed in [\[44\]](#) provided with [\[45\]](#), and it is possible that the additional difficulty is even more substantial for us since we need to nail down the exact threshold. We refer the reader to the discussions at the beginning of Sections [3.2](#) and [3.3](#) for overviews of the proofs for [\(3.2\)](#) and [\(3.3\)](#), respectively.

3.1.3 Notations

We record in this subsection a list of notations that we shall use throughout the paper. First recall that we have two vertex sets V, \mathbf{V} with $|V| = |\mathbf{V}| = n$, and \mathbf{P}, \mathbf{Q} are two probability measures on pairs of random graphs on V and \mathbf{V} defined previously. In addition, for an edge $e \in E_0$, G_e denotes for the indicator of the event that e is an edge in G , and the similar notation applies for $G_{\mathbf{e}}$ with $\mathbf{e} \in E_0$. The following is a collection of notational conventions we shall follow.

- $B(A, \mathbf{A})$, $B(V, \mathbf{V})$ and $B(V, \mathbf{V}, A, \sigma)$. For any two sets $A \subset V$ and $\mathbf{A} \subset \mathbf{V}$, we denote

$B(A, A)$ for the set of embeddings from A to A . In particular, $B(V, V)$ is the set of bijections from V to V . For any subset $A \subset V$ and any embedding $\sigma : A \rightarrow V$, let $B(V, V, A, \sigma) \subset B(V, V)$ be the set of bijections which inhibit to A as σ .

- *Induced subgraphs H_A and H^A .* For a graph $H = (V, E)$ and a subset $A \subset V$, define $H_A = (A, E_A)$ to be the induced subgraph of H in A , and $H^A = (V, E^A)$ to be the subgraph of H obtained by deleting all edges within A . Similar notations $H_A = (A, E_A), H^A = (V, E^A)$ apply for any graph $H = (V, E)$ and any subset $A \subset V$.

- *The probability measure \mathcal{Q} .* Define a probability measure \mathcal{Q} on the space of triples

$$\Omega = \{(\pi^*, G, \mathbb{G}) : \pi^* \in B(V, V), G, \mathbb{G} \text{ are subgraphs of } (V, E_0), (V, E_0)\} \quad (3.5)$$

as follow: the marginal distributions of π^* under \mathcal{Q} is uniform on $B(V, V)$; conditioned on π^* , (G, \mathbb{G}) is a pair of correlated Erdős-Rényi graphs given as in (3.1). It is clear that \mathcal{Q} is nothing but the marginal distribution for the last two coordinates of \mathcal{Q} .

- *Edge bijection and permutation.* For any bijection $\pi \in B(V, V)$ (respectively permutation ϕ on V), define the bijection between E_0 and E_0 (respectively permutation on E_0) induced by π (respectively ϕ) as Π (respectively Φ). That is to say, for any $(u, v) \in E_0$, we have $\Pi((u, v)) = (\pi(u), \pi(v))$ (respectively $\Phi((u, v)) = (\phi(u), \phi(v))$).

- *Edge orbits and \mathcal{O}_π .* Assume $\pi^* \in B(V, V)$ is fixed. For any $\pi \in B(V, V)$, $\phi \stackrel{\text{def}}{=} \pi^{-1} \circ \pi^*$ is a permutation on V . The induced edge permutation Φ on E_0 decomposes E_0 into disjoint edge cycles. We call these cycles as edge orbits induced by π and denote \mathcal{O}_π for the collection of all such edge orbits. The point of this notation is that, the families of random variables $\{(G_e, \mathbb{G}_{\Pi(e)}) : e \in O\}$ are mutually independent under the law $\mathcal{Q}[\cdot \mid \pi^*]$ for distinct edge orbits $O \in \mathcal{O}_\pi$.

- *π -intersection graphs $\mathcal{H}_\pi = (V, \mathcal{E}_\pi)$.* For any triple $(\pi, G, \mathbb{G}) \in \Omega$ defined as in (3.5), the π -intersection graph $\mathcal{H}_\pi = (V, \mathcal{E}_\pi)$ of (G, \mathbb{G}) is a subgraph of (V, E_0) , where $(u, v) \in \mathcal{E}_\pi$ if and only if (u, v) is an edge in G and $(\pi(u), \pi(v))$ is also an edge in \mathbb{G} .

Acknowledgement. We warmly thank Nicholas Wormald, Yihong Wu and Jiaming Xu

for stimulating discussions. Hang Du is partially supported by the elite undergraduate training program of School of Mathematical Sciences in Peking University.

3.2 Possibility for partial recovery

In this section, we prove Theorem 3.1-(3.2), where $\lambda \geq \lambda_* + \varepsilon$ for some arbitrary and fixed $\varepsilon > 0$. The construction of the estimator $\hat{\pi}$ naturally takes inspiration from the detection statistics as in [12, Section 2]. For instance, we may simply define $\hat{\pi}$ to be the maximizer for $\max_{\pi} \max_{U: |U| \geq c_{\lambda} n} \frac{|\mathcal{E}_{\pi}(U)|}{|U|}$ since this was shown in [12] as an efficient statistics for testing correlation against independence (or alternatively for convenience of analysis, choose $\hat{\pi}$ as an arbitrary matching whose intersection graph has maximal subgraph density exceeding a certain threshold). While this estimator may in fact has non-vanishing overlap with the true matching π^* , it seems rather difficult to prove as we now explain. In order to justify the estimator one has to show that typically any matching that has vanishing overlap with π^* can not be a maximizer, and claims of this type are usually proved via a first moment computation. In [12] a first moment computation along this line was carried out to show that when two graphs are independent there is no matching whose intersection graph has subgraph with at least $c_{\lambda} n$ vertices and also a large edge density. In the independent case, all matchings are symmetric with each other (since graphs are generated independently with π^*) and thus the first moment computation is rather straightforward. However, in this paper the computation needs to be carried out for correlated graphs and thus overlapping structures with π^* for different matchings play a significant role such that not only they complicate the computation but also it seems they actually will lead to a blowup of the first moment (but this does not necessarily imply that “bad” things do happen since the blowup may come from an event of small probability). One common approach in this case is to introduce further truncation, and a natural truncation that comes to mind is to pose an upper bound on the maximal subgraph edge density, since a plausible way for the first moment to blow up is due to an event of small probability where the maximal subgraph

edge density on an intersection graph is excessively high. With these intuitions in mind, we define our estimator $\hat{\pi} = \hat{\pi}(G, \mathbf{G})$ as follows.

Definition 3.1. Fix some $0 < \eta < \frac{\varrho - \alpha^{-1}}{4}$, where ϱ is short for $\varrho(\lambda)$. For any matching $\pi \in B(V, V)$, we say it is a reasonable candidate of π^* if its π -intersection graph \mathcal{H}_π of (G, \mathbf{G}) satisfies the following two conditions:

- (i) The edge-vertex ratio of any nonempty subgraph of \mathcal{H}_π does not exceed $\varrho + \eta$;
- (ii) There is a subgraph of \mathcal{H}_π with size at least $c_\lambda n$ and edge-vertex ratio at least $\varrho - \eta$.

If the set of reasonable candidates is nonempty, choose one of them as $\hat{\pi}$ arbitrarily. Otherwise pick a $\hat{\pi} \in B(V, V)$ randomly.

Note that $\mathcal{H}_{\pi^*} \sim \mathbf{G}(n, \frac{\lambda}{n})$ and thus by Proposition 3.2 we see that π^* is a reasonable candidate with probability tending to 1 as $n \rightarrow \infty$. Therefore, in order to prove (3.2) it suffices to prove the following proposition.

Proposition 3.4. There exists $\delta = \delta(\alpha, \varepsilon) > 0$ such that the following holds. Denote \mathcal{B} for the event that there exists a reasonable candidate π with $\text{overlap}(\pi^*, \pi) \leq \delta n$. Then $\mathcal{Q}[\mathcal{B}] \rightarrow 0$ as $n \rightarrow \infty$.

We hope to bound $\mathcal{Q}[\mathcal{B}]$ by the first moment method and we hope that Condition (i) for reasonable candidate will help controlling the moment. It was not *a priori* clear why Condition (i) suffices, and in fact even after completing the proof we do not feel that there is a one-sentence explanation on why Condition (i) suffices since the proof of Proposition 3.4 seems to involve fairly nontrivial probability and combinatorics.

A key estimate required to prove Proposition 3.4 is the tail probability for \mathcal{H}_π to satisfy Condition (ii). Due to complication arising from correlations, this is not a very straightforward computation since on the one hand we only have independence between different edge orbits (see Section 3.2.1 for definition) and on the other hand orbits with different sizes have different large deviation rates. This motivates us to classify π according

to the structure of orbits (as in Section 3.2.1) and then perform a union bound over matchings via a union bound over different classes of matchings (as in Section 3.2.2). Along the way, we will postpone proofs for a few technical lemmas/propositions into appendices to maintain a smooth flow of presentation.

3.2.1 Orbits and tail probabilities

In this subsection, for convenience we consider π^* as fixed. Mathematically speaking, we condition on the realization of π^* and we slightly abuse the notation by denoting π^* as its realization. Fix some $A \subset V$. For any $\pi \in B(V, V)$, let $\phi \stackrel{\text{def}}{=} \pi^{-1} \circ \pi^*$ and recall the definition of Π and Φ in Section 3.1.3. Similar with \mathcal{O}_π in Section 3.1.3, we also define a set of edge orbits $(\mathcal{O}_\pi)_A$ in $(E_0)_A$ as the collection of orbits induced by the mapping Φ . More precisely, each orbit has the form (e_1, \dots, e_k) with $e_1, \dots, e_k \in (E_0)_A, e_{i+1} = \Phi(e_i)$ for $1 \leq i \leq k-1$ and in addition satisfies

$$e_1 = \Phi(e_k), \quad \text{or} \quad \Phi^{-1}(e_1) \notin (E_0)_A, \Phi(e_k) \notin (E_0)_A. \quad (3.6)$$

It is clear that $(\mathcal{O}_\pi)_A$ is a partition of the edge set $(E_0)_A$. Again, the virtue of decomposing $(E_0)_A$ into orbits in $(\mathcal{O}_\pi)_A$ is that the families of random variables $\{(G_e, \mathbf{G}_{\Pi(e)}) : e \in O\}$ are mutually independent under $\mathcal{Q}[\cdot \mid \pi^*]$ for distinct $O \in (\mathcal{O}_\pi)_A$.

From the definition we see that any orbit $O \in (\mathcal{O}_\pi)_A$ is a cycle (if the former occurs in (3.6)) or a chain (if the latter occurs in (3.6)), and we shall call O a k -cycle (respectively k -chain) if it is a cycle (respectively chain) with length k (i.e., with k edges in the orbit). For convenience, we denote by $\text{LCM}(x, y)$ as the least common multiple for two integers x and y . The following lemma characterizes the structure of orbits in $(\mathcal{O}_\pi)_A$ in a more detailed way. The lemma is relatively obvious and an illustrative explanation can be found in [45, Subsection 5.1]. As a result, we omit its proof.

Lemma 3.5. *For an edge $(u, v) \in (E_0)_A$, the following hold:*

(a) *The orbit of (u, v) in $(\mathcal{O}_\pi)_A$ is a cycle if and only if the node cycle of u with respect to*

ϕ is entirely contained in A and so is that for v .

(b) If the node cycle of u with respect to ϕ are disjoint from that of v and both cycles are entirely contained in A with lengths x and y respectively, then the orbit $O \in (\mathcal{O}_\pi)_A$ containing (u, v) is a $\text{LCM}(x, y)$ -cycle.

(c) If u, v are in the same node cycle with respect to ϕ and if this cycle is entirely contained in A with length x , then the orbit $O \in (\mathcal{O}_\pi)_A$ containing (u, v) is an x -cycle or an $\frac{x}{2}$ -cycle, with the latter happens only when x is even and $v = \phi^{x/2}(u)$. The cycles in the latter case are called special.

For $\pi \in B(V, V)$ and any edge orbit $O \in (\mathcal{O}_\pi)_A$, let $\mathcal{E}_O = O \cap \mathcal{H}_\pi$. We have the following exponential moments for $|\mathcal{E}_O|$ under the law $\mathcal{Q}[\cdot \mid \pi^*]$. We write $a_n \ll b_n$ if $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$.

Proposition 3.6. *For any θ with $1 \ll e^\theta \ll n$, the following hold with $\mu_1 = 1 + e^\theta/n^{1+\alpha+o(1)}$ and $\mu_2 = (\lambda + o(1))e^\theta/n$:*

- For any k -cycle $O_k \in (\mathcal{O}_\pi)_A$,

$$\mathbb{E}_{(G, \mathcal{G}) \sim \mathcal{Q}[\cdot \mid \pi^*]} \exp(\theta |\mathcal{E}_{O_k}|) = \mu_1^k + \mu_2^k. \quad (3.7)$$

- For any k -chain $O_k \in (\mathcal{O}_\pi)_A$,

$$\mathbb{E}_{(G, \mathcal{G}) \sim \mathcal{Q}[\cdot \mid \pi^*]} \exp(\theta |\mathcal{E}_{O_k}|) \leq \mu_1^k + e^\theta n^{-1-2\alpha+o(1)} \mu_2^k. \quad (3.8)$$

At this moment, we need to give a “cutoff” between short cycles and long cycles. To this end, we define

$$N = \begin{cases} \lfloor (1 - \alpha)^{-1} \rfloor, & \alpha < 1, \\ \lfloor (\varrho - \eta - 1)^{-1} \rfloor + 1, & \alpha = 1. \end{cases} \quad (3.9)$$

(We keep in mind that $\leq N$ means short and $> N$ means long.) For a fixed $\pi \in B(V, V)$, let E_s be the total number of edges in \mathcal{H}_π coming from special cycles in $(\mathcal{O}_\pi)_A$, and for $1 \leq k \leq N$ let E_k be the total number of edges in \mathcal{H}_π coming from non-special k -cycles

in $(\mathcal{O}_\pi)_A$, and let E_{N+1} be the total number of edges in \mathcal{H}_π coming from chains and non-special cycles in $(\mathcal{O}_\pi)_A$ with lengths at least $N+1$. Proposition 3.6 leads to the following estimates on tail probabilities.

Lemma 3.7. *Let $\alpha_k = \frac{k-1}{k}$ for $1 \leq k \leq N$ and let $\alpha_{N+1} = \alpha \wedge \frac{N}{N+1}$. For some $M = o(n \log n)$ and any $0 \leq x \leq 2\varrho n$, the following hold:*

$$\mathcal{Q}[E_s \geq x \mid \pi^*] \leq \exp(M - x \log n), \quad (3.10)$$

$$\mathcal{Q}[E_k \geq x \mid \pi^*] \leq \exp(M - \alpha_k x \log n) \text{ for } 1 \leq k \leq N, \quad (3.11)$$

$$\mathcal{Q}[E_{N+1} \geq x \mid \pi^*] \leq \exp(M - \alpha_{N+1} x \log n). \quad (3.12)$$

Proof. Suppose there are S_k special k -cycles, L_k non-special k -cycles and T_k many k -chains in $(\mathcal{O}_\pi)_A$. Then $\sum_{k \geq 1} k S_k \leq n$ since each vertex belongs to no more than one special cycle, and $\sum_{k \geq 1} k(L_k + T_k) \leq n^2$ since the total number of edges in $(\mathcal{O}_\pi)_A$ is bounded by n^2 .

By taking $\theta = \log n - \log \log n$ and applying Markov's inequality, the left hand side of (3.10) is bounded by

$$\begin{aligned} & e^{-\theta x} \prod_{k \geq 1} (\exp(\theta |\mathcal{E}_{O_k}|))^{S_k} \stackrel{(3.7)}{=} e^{-\theta x} \prod_{k \geq 1} (\mu_1^k + \mu_2^k)^{S_k} \leq e^{-\theta x} (\mu_1 + \mu_2)^{\sum_{k \geq 1} k S_k} \\ & \leq e^{-\theta x} \left[1 + \frac{e^\theta}{n^{1+\alpha+o(1)}} + \frac{(\lambda + o(1))e^\theta}{n} \right]^n \leq \exp(o(n) - x \log n + x \log \log n). \end{aligned} \quad (3.13)$$

Similarly, by taking $\theta = \alpha_k \log n - \log \lambda$ and applying Markov's inequality together with (3.7) again, we see the left hand side of (3.11) is bounded by

$$\begin{aligned} & e^{-\theta x} [\mu_1^k + \mu_2^k]^{L_k} \leq e^{-\theta x} \left(1 + \frac{e^\theta}{n^{1+\alpha+o(1)}} + \frac{(\lambda^k + o(1))e^{k\theta}}{n^k} \right)^{n^2} \\ & \leq \exp(x \log \lambda + n + o(n) - \alpha_k x \log n). \end{aligned} \quad (3.14)$$

Finally, applying Markov's inequality together with (3.7) and (3.8), we see for any θ

with $1 \ll e^\theta \ll n$, the left hand side of (3.12) is bounded by

$$\begin{aligned}
& e^{-\theta x} \prod_{k \geq N+1} \left(\mu_1^k + \mu_2^k \right)^{L_k} \prod_{k \geq 1} \left(\mu_1^k + e^\theta n^{-1-2\alpha+o(1)} \mu_2^k \right)^{T_k} \\
& \leq e^{-\theta x} \prod_{k \geq N+1} \left(\mu_1^{N+1} + \mu_2^{N+1} \right)^{\frac{kL_k}{N+1}} \prod_{k \geq 1} \mu_1^{kT_k} (1 + e^\theta n^{-1-2\alpha+o(1)} \mu_2^k)^{T_k} \\
& \leq e^{-\theta x} \left(1 + e^\theta n^{-1-\alpha+o(1)} + (\lambda^{N+1} + o(1)) e^{(N+1)\theta} n^{-N-1} \right)^{n^2} \left(1 + e^{2\theta} n^{-2-2\alpha+o(1)} \right)^{n^2}.
\end{aligned} \tag{3.15}$$

When $\alpha < 1$, we pick θ such that $e^\theta n^{-1-\alpha+o(1)}$ in the first bracket above equals to n^{-1} , then (3.15) becomes $\exp(n + o(n) - (\alpha - o(1))x \log n)$, and here crucially we used that $N + 1 > (1 - \alpha)^{-1}$, implying $n^{(N+1)(\alpha-o(1))-N-1} \ll n^{-1}$. When $\alpha = 1$, we just pick $\theta = \alpha_{N+1} \log n - \log \lambda$ and (3.15) becomes $\exp(x \log \lambda + n + o(n) - \alpha_{N+1} x \log n)$.

Take $M = (2\varrho + 1)(n \log \log n + o(1)n \log n) = o(n \log n)$ (with a suitable choice of $o(1)$ originating from $o(1)$ -terms as above). Then with all of the aforementioned bounds, (3.10), (3.11) and (3.12) hold for large n , as desired. \square

3.2.2 Proof of Proposition 3.4

For $A \subset V$ and any $\sigma \in B(A, \mathbb{V})$, choose some $\bar{\sigma} \in B(V, \mathbb{V}, A, \sigma)$ as an extension of σ on V . It is easy to see that the set $(\mathcal{O}_{\bar{\sigma}})_A$ does not depend on the choice of extension $\bar{\sigma}$, and hence the set of node cycles of $\phi = \bar{\sigma}^{-1} \circ \pi^*$ which are entirely contained in A is also well-defined. For any sequence of non-negative integers n_1, \dots, n_N satisfying $\sum_{k=1}^N k n_k \leq |A|$, we define $S(A, n_1, \dots, n_N) \subset B(A, \mathbb{V})$ to be

$$\{\sigma \in B(A, \mathbb{V}) : \phi = \bar{\sigma}^{-1} \circ \pi^* \text{ has } n_k \text{ node cycles with length } k \text{ in } A \text{ for } 1 \leq k \leq N\}.$$

Lemma 3.8. *For $A \subset V$ with $|A| = T$ and non-negative integers n_1, \dots, n_N satisfying $\sum_{k=1}^N k n_k \leq T$, it holds that*

$$|S(A, n_1, \dots, n_N)| \leq \frac{n(n-1) \cdots (n-T+1)}{\prod_{k=1}^N k^{n_k} n_k!} = \exp(O(n) + (T - n_1 - \cdots - n_N) \log n).$$

Proof. Define

$$S = \bigcup_{\sigma \in S(A, n_1, \dots, n_N)} B(V, V, A, \sigma)$$

to be the collection for all extensions $\pi \in B(V, V)$ of some $\sigma \in S(A, n_1, \dots, n_N)$. This is a disjoint union since for π_1, π_2 from different $B(V, V, A, \sigma_1), B(V, V, A, \sigma_2)$, we have $\pi_1|_A = \sigma_1 \neq \sigma_2 = \pi_2|_A$, implying $\pi_1 \neq \pi_2$ (here we denote by $\pi|_A$ the restriction of π on A). In addition, note that for any $\pi \in S$, $\phi = \pi^{-1} \circ \pi^*$ contains at least n_k many k -node cycles for all $1 \leq k \leq N$. By [4, Theorem 1] (see also [44, Lemma 13]), the number for such ϕ is no more than $\frac{n!}{\prod_{k=1}^N k^{n_k} n_k!}$. As a result,

$$|S| = (n - T)! \cdot |S(A, n_1, \dots, n_N)| \leq \frac{n!}{\prod_{k=1}^N k^{n_k} n_k!}, \quad (3.16)$$

which yields the desired inequality (in the lemma statement) and the equality follows from Stirling's formula. \square

Proof of Proposition 3.4. For any $A \subset V$ with $|A| = T \in [c_\lambda n, n]$ and any embedding $\sigma \in B(A, V)$, denote \mathcal{B}_σ for the event that the $\bar{\sigma}$ -intersection graph $\mathcal{H}_{\bar{\sigma}}$ satisfies the following:

- (i) $|\mathcal{E}_{\bar{\sigma}}(A)| \geq (\varrho - \eta)T$;
- (ii) $|\mathcal{E}_{\bar{\sigma}}(U)| \leq (\varrho + \eta)|U|$ for any $U \subset A$.

Again, we note that \mathcal{B}_σ does not depend on the choice of extension $\bar{\sigma}$.

For each $\pi^* \in B(V, V)$, it is clear that conditioned on π^* , \mathcal{B} implies \mathcal{B}_σ happens for some A with size at least $c_\lambda n$ and some $\sigma \in B(A, V)$ which agrees with π^* on less than δn vertices in A . Then a simple union bound yields that

$$\begin{aligned} \mathcal{Q}[\mathcal{B} \mid \pi^*] &\leq \sum_{T \geq c_\lambda n} \sum_{|A|=T} \sum_{n_1, \dots, n_N} \sum_{\sigma \in S(A, n_1, \dots, n_N)} \mathcal{Q}[\mathcal{B}_\sigma \mid \pi^*] \\ &\leq \sum_{|T| \geq c_\lambda n} \sum_{|A|=T} \sum_{n_1, \dots, n_N} |S(A, n_1, \dots, n_N)| \times \sup_{\sigma \in S(A, n_1, \dots, n_N)} \mathcal{Q}[\mathcal{B}_\sigma \mid \pi^*], \end{aligned} \quad (3.17)$$

where the summation for n_i 's is over all non-negative integers n_1, \dots, n_N with $\sum_{k=1}^N k n_k \leq T$ and $n_1 \leq \delta n$ (since σ overlaps on less than δn vertices with π^*).

We next turn to bound $\sup_{\sigma \in S(A, n_1, \dots, n_N)} \mathcal{Q}[\mathcal{B}_\sigma \mid \pi^*]$. For any fixed T, A, n_1, \dots, n_N and $\sigma \in S(A, n_1, \dots, n_N)$, let $A_i \subset A$ be the set of vertices in node cycles of $\phi = \bar{\sigma}^{-1} \circ \pi^*$ with length i in A . Then $|A_i| = in_i$ since there are n_i many i -cycles in A . Let x_{ij} be the number of edges in \mathcal{H}_π with two end points in A_i and A_j respectively, but not in special cycles. Then by Lemma 3.5,

$$E_k = \sum_{i,j: \text{LCM}(i,j)=k} x_{ij}, \text{ for all } 1 \leq k \leq N.$$

Now from (ii) in \mathcal{B}_σ , we get for all $1 \leq m \leq N$,

$$\sum_{k=1}^m E_k = \sum_{i,j: \text{LCM}(i,j) \leq m} x_{ij} \leq |\mathcal{E}_{\bar{\sigma}}(A_1 \cup A_2 \cup \dots \cup A_m)| \leq (\varrho + \eta) \sum_{k=1}^m kn_k. \quad (3.18)$$

Combined with (i) in \mathcal{B}_σ , this motivates us to define

$$\begin{aligned} \Sigma_T &= \left\{ (x_0, \dots, x_{N+1}) \in \mathbb{Z}^{N+2} : 0 \leq x_0, \dots, x_{N+1} \leq \varrho T, \sum_{k=0}^{N+1} x_k \geq (\varrho - \eta)T \right\}, \\ \Delta_{n_1, \dots, n_N} &= \left\{ (x_0, \dots, x_{N+1}) \in \mathbb{Z}^{N+2} : \sum_{k=1}^m x_k \leq (\varrho + \eta) \sum_{k=1}^m kn_k, \forall 1 \leq m \leq N \right\}. \end{aligned}$$

Somewhat mysteriously, the restriction to Δ_{n_1, \dots, n_N} which originates from our further truncation as in (ii) of \mathcal{B}_σ suffices to control the (otherwise) blowup of the first moment. The seemingly computational coincidence is encapsulated in the following (purely algebraic) lemma. Denote $M(T, n_1, \dots, n_N)$ as

$$\min \left\{ \sum_{k=1}^N n_k - T + x_0 + \sum_{k=1}^{N+1} \alpha_k x_k \mid (x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N} \right\}. \quad (3.19)$$

In light of Lemmas 3.7 and 3.8, $M(T, n_1, \dots, n_N)$ captures the tradeoff between probability and enumeration.

Lemma 3.9. *There exist constants $\delta, \delta_0 > 0$, such that for any $T \in [c_\lambda n, n]$ and n_1, \dots, n_N satisfying $n_1 \leq \delta n$ and $n_1 + 2n_2 + \dots + Nn_N \leq T$, we have that $M(T, n_1, \dots, n_N) \geq \delta_0 T$.*

With Lemma 3.9 at hand, we see the supremum of $\mathcal{Q}[\mathcal{B}_\sigma \mid \pi^*]$ for $\sigma \in \mathcal{S}(A, n_1, \dots, n_N)$ is bounded by

$$\begin{aligned}
& \sum_{(x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N}} \mathcal{Q}[E_s \geq x_0 \text{ and } E_k \geq x_k, 1 \leq k \leq N \mid \pi^*] \\
& \stackrel{\text{independence}}{=} \sum_{(x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N}} \mathcal{Q}[E_s \geq x_0 \mid \pi^*] \prod_{k=1}^{N+1} \mathcal{Q}[E_k \geq x_k \mid \pi^*] \\
& \stackrel{\text{Lemma 3.7}}{\leq} \sum_{(x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N}} \exp \left(o(n \log n) - \left[x_0 + \sum_{k=1}^{N+1} \alpha_k x_k \right] \log n \right). \quad (3.20)
\end{aligned}$$

Plugging (3.20) and Lemma 3.8 into (3.17), we see $\mathcal{Q}[\mathcal{B} \mid \pi^*]$ is bounded by

$$\begin{aligned}
& \sum_{T \geq c_\lambda n} \sum_{|A|=T} \sum_{n_1, \dots, n_N} \sum_{(x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N}} \\
& \quad \exp \left(o(n \log n) - \left[\sum_{k=1}^N n_k - T + x_0 + \sum_{k=1}^{N+1} \alpha_k x_k \right] \log n \right) \\
& \leq \sum_{T \geq c_\lambda n} \binom{n}{T} T^N (\varrho n)^{N+2} \exp(o(n \log n) - M(T, n_1, \dots, n_N) \log n). \quad (3.21)
\end{aligned}$$

By Lemma 3.9, we see (3.21) is no more than

$$\sum_{T \geq c_\lambda n} \exp(o(n \log n) - \delta_0 T \log n) = o(1).$$

Thus, $\mathcal{Q}[\mathcal{B} \mid \pi^*] = o(1)$. Since this is invariant for any $\pi^* \in B(V, V)$, we complete the proof of (3.2). \square

3.3 Impossibility for partial recovery

In this section we prove Theorem 3.1-(3.3). Recall that we are now in the regime of $\lambda \leq \lambda_* - \varepsilon$ for some arbitrary and fixed $\varepsilon > 0$. At the first glance, this might seem trivial (as the authors have wrongly speculated) in light of Part (ii) in Proposition 3.3. However, it turns out to be not at all obvious how to derive impossibility for correctly matching a positive fraction of vertices from impossibility of detection, since for instance we may

correctly match a linear sized independent set in \mathcal{H}_{π^*} but the intersection graph for this matching can be similar to that of a typical matching for independent graphs (and thus has no power for detection).

Fix an arbitrary $\delta > 0$. For a pair of graphs (G, \mathbb{G}) , denote $\mathcal{Q}_{G, \mathbb{G}}$ for the posterior distribution of π^* under the law \mathcal{Q} when (G, \mathbb{G}) are given as observations. If there exists an estimator $\hat{\pi} = \hat{\pi}(G, \mathbb{G})$ such that $\hat{\pi}$ correctly matches at least a δ -fraction of vertices with non-vanishing probability, then $\mathcal{Q}_{G, \mathbb{G}}$ must be somehow concentrate around $\hat{\pi}$ with non-vanishing probability. In light of this, for any $\tilde{\pi} \in B(V, V)$, we define

$$M(G, \mathbb{G}, \tilde{\pi}) = \mathcal{Q}_{G, \mathbb{G}}[\text{overlap}(\pi^*, \tilde{\pi}) \geq \delta n] = \frac{1}{\mathcal{Q}[G, \mathbb{G}]} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \quad (3.22)$$

to be the measure of $\mathcal{Q}_{G, \mathbb{G}}$ on the set of $\pi \in B(V, V)$ which agrees with $\tilde{\pi}$ on more than δn vertices. Further, we set

$$W(G, \mathbb{G}) = \max_{\tilde{\pi} \in B(V, V)} M(G, \mathbb{G}, \tilde{\pi}). \quad (3.23)$$

Then the following proposition is the key to the proof of (3.3).

Proposition 3.10. *For any $\delta > 0$, we have $\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} W(G, \mathbb{G}) \rightarrow 0$ as $n \rightarrow \infty$.*

Proof of Theorem 3.1-(3.3) assuming Proposition 3.10. For any estimator $\hat{\pi} = \hat{\pi}(G, \mathbb{G})$ (perhaps with additional randomness), denote $\mathbb{P}_{G, \mathbb{G}}$ for the law of $\hat{\pi}$ given (G, \mathbb{G}) . Then the probability that $\text{overlap}(\pi^*, \hat{\pi}) \geq \delta n$ can be expressed as

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} \mathbb{E}_{\hat{\pi} \sim \mathbb{P}_{G, \mathbb{G}}} \mathcal{Q}_{G, \mathbb{G}}[\text{overlap}(\pi^*, \hat{\pi}) \geq \delta n] \leq \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} W(G, \mathbb{G}),$$

which is $o(1)$ by Proposition 3.10. This completes the proof. \square

The starting point for the proof of Proposition 3.10 is the observation that for any $K \leq \delta n$ overlapping on more than δn vertices necessarily implies overlapping on some $A \subset V$ with $|A| = K$, and indeed we will choose $K = n^\beta$ for β close to 1 but less than 1. On the one hand, we choose β close to 1 to avoid losing too much information; on the other hand, we choose β strictly less than 1 since it seems second moment computation applies

better when K is smaller. The second moment computation also leads to the next very simple but useful observation that $\max_{x \in S} x^2 \leq \sum_{x \in S} x^2$, which then allows us to upper-bound the maximum of posterior probability by its second moment. At the first glance, this bound seems really loose to be useful, but it turns out that for $K \geq n^\beta$ with some β very close to 1, applying this simple inequality does yield an efficient upper bound on the probability of correctly matching vertices in A (see (3.41)). Therefore, the main technical obstacle is now reduced to a second moment computation (see Proposition 3.15) for which we take inspiration from [12] and we remark here that a few additional truncations (on top of those in [12]) are necessary in order for our purpose. Finally, we note that in light of Proposition 3.3, we can carry out many computations under the measure \mathbf{P} for independent graphs, which in many ways simplifies our analysis (see e.g., Proposition 3.11 and Proof of Proposition 3.10 assuming Proposition 3.11).

3.3.1 Truncations for the second moment

As mentioned earlier, we necessarily need to introduce truncations in order to prevent the second moment from blowing up. This truncation is expressed as some “good” event \mathcal{G} as in Definition 3.2 below, and as we will see \mathcal{G} is measurable with respect to (π^*, G, \mathbb{G}) and satisfies $\mathcal{Q}[\mathcal{G}] \rightarrow 1$ as $n \rightarrow \infty$. Thus, it would be useful to reduce Proposition 3.10 to a version with truncation as follows.

Proposition 3.11. *For any $\delta > 0$, we have*

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbf{P}} \frac{1}{\mathbf{P}[G, \mathbb{G}]} \max_{\tilde{\pi} \in B(V, V)} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (3.24)$$

Proof of Proposition 3.10 assuming Proposition 3.11. For any $\tilde{\pi} \in B(V, V)$, we can upper-bound $M(G, \mathbb{G}, \tilde{\pi})$ by

$$\frac{1}{\mathbf{Q}[G, \mathbb{G}]} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} + \frac{\mathcal{Q}[\mathcal{G}^c, G, \mathbb{G}]}{\mathbf{Q}[G, \mathbb{G}]},$$

where $\mathcal{Q}[\mathcal{G}^c, G, \mathbb{G}] = \sum_{\pi} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}^c}(\pi, G, \mathbb{G})$. This yields that

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} W(G, \mathbb{G}) \leq \frac{1}{\mathcal{Q}[G, \mathbb{G}]} \max_{\tilde{\pi} \in B(V, V)} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} + \mathcal{Q}[\mathcal{G}^c]. \quad (3.25)$$

Recall the definition of \mathbf{P} and the fact that $\text{TV}(\mathbf{P}, \mathbf{Q}) = o(1)$ given in Proposition 3.3-(ii).

Define the event \mathcal{G}_0 as

$$\mathcal{G}_0 = \left\{ (G, \mathbb{G}) : \frac{\mathcal{Q}[G, \mathbb{G}]}{\mathbf{P}[G, \mathbb{G}]} \geq \frac{1}{2} \right\}.$$

Since $(G, \mathbb{G}) \in \mathcal{G}_0^c$ implies $\mathcal{Q}[G, \mathbb{G}] \leq \mathbf{P}[G, \mathbb{G}] - \mathcal{Q}[G, \mathbb{G}]$, we have that

$$\mathcal{Q}[\mathcal{G}_0^c] \leq \int_{\mathcal{G}_0^c} (\text{d}\mathbf{P} - \text{d}\mathbf{Q}) \leq \text{TV}(\mathbf{P}, \mathbf{Q}) = o(1). \quad (3.26)$$

Since $\frac{1}{\mathcal{Q}[G, \mathbb{G}]} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \leq 1$ holds for any $\tilde{\pi} \in B(V, V)$, we can upper-bound the right hand side of (3.25) by

$$\begin{aligned} & \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} \frac{\mathbf{1}_{\mathcal{G}_0}(G, \mathbb{G})}{\mathcal{Q}[G, \mathbb{G}]} \max_{\tilde{\pi} \in B(V, V)} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} + \mathcal{Q}[\mathcal{G}_0^c] + \mathcal{Q}[\mathcal{G}^c] \\ & \leq \mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}} \frac{2}{\mathbf{P}[G, \mathbb{G}]} \max_{\tilde{\pi} \in B(V, V)} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} + \mathcal{Q}[\mathcal{G}_0^c] + \mathcal{Q}[\mathcal{G}^c] \\ & \leq \mathbb{E}_{(G, \mathbb{G}) \sim \mathbf{P}} \frac{2}{\mathbf{P}[G, \mathbb{G}]} \max_{\tilde{\pi} \in B(V, V)} \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} + 2 \text{TV}(\mathbf{P}, \mathbf{Q}) + \mathcal{Q}[\mathcal{G}_0^c] + \mathcal{Q}[\mathcal{G}^c], \end{aligned}$$

which is $o(1)$ by Proposition 3.11, (3.26) and the choice of \mathcal{G} . This concludes the proof. \square

Next, we give the precise definition of the good event \mathcal{G} and further introduce the concept of good set. These definitions may be a bit perplexing at the first glance, but the motivations behind all the constraints would become clear in Section 3.4.3 (a reader may skip this definition for now and reference back when certain constraints are used in controlling moments in later proofs). We first introduce some constants $\xi, \zeta, \beta, C, \delta_1$.

Denote

$$\xi = \frac{\varrho + \alpha^{-1}}{2}, \quad (3.27)$$

where ϱ is short for $\varrho(\lambda)$. Note that under the assumption $\lambda \leq \lambda^* - \varepsilon$, we have $\xi < \alpha^{-1}$.

Since $\alpha < 1$, it is possible to take a constant $\zeta > 1$ such that

$$1 + \zeta(\alpha - 1) < 2 - \zeta. \quad (3.28)$$

Then, we choose β such that

$$(1 - \alpha) \vee \frac{1 + \zeta(\alpha - 1)}{2 - \zeta} < \beta < 1, \quad (3.29)$$

and some large integer C such that

$$\alpha(\xi + C^{-1}) < 1. \quad (3.30)$$

Finally, we pick $\delta_1 > 0$ such that

$$\delta_1 < (1 - \alpha\xi) \wedge C^{-1}\beta. \quad (3.31)$$

Definition 3.2. For a graph $\mathcal{H} = (V, \mathcal{E})$, we say \mathcal{H} is admissible if \mathcal{H} satisfies the following properties:

- (i) For any $A \subset V$, it holds $|\mathcal{E}(A)| \leq \xi|A|$;
- (ii) For any $A \subset V$ satisfying $|A| \leq n/\log n$, it holds $|\mathcal{E}(A)| \leq \zeta|A|$;
- (iii) The maximal degree of \mathcal{H} is less than $\log n$;
- (iv) Any connected subgraph of \mathcal{H} with size less than $\log \log n$ contains at most one cycle;
- (v) For any $k \geq 3$, the number of cycles with length k in \mathcal{H} is bounded by $n^{\delta_1 k}$.

Let \mathcal{G} be the event that \mathcal{H}_{π^*} , the π^* -intersection graph of G and \mathbf{G} , is admissible.

It is clear that \mathcal{G} is measurable with respect to the triple (π^*, G, \mathbf{G}) . We next show that \mathcal{G} is indeed a typical event under \mathcal{Q} .

Lemma 3.12. $\mathcal{Q}[\mathcal{G}] \rightarrow 1$ as $n \rightarrow \infty$.

Proof. Note that \mathcal{H}_{π^*} has the law of an Erdős-Rényi graph $\mathbf{G}(n, \frac{\lambda}{n})$, it suffice to bound the probability that either of (i)-(v) fails for such an Erdős-Rényi graph. $\mathbb{P}[(i) \text{ fails}] = o(1)$

follows from Proposition 3.2. For (ii), since $\zeta > 1$, we can take a union bound as (denoting by $\mathbf{B}(m, q)$ as a binomial variable with m trials and success probability q)

$$\begin{aligned} \mathbb{P}[(ii) \text{ fails}] &\leq \sum_{k \leq n/\log n} \binom{n}{k} \mathbb{P} \left[\mathbf{B} \left(\binom{k}{2}, \frac{\lambda}{n} \right) > \zeta k \right] \\ &\leq \sum_{k \leq n/\log n} \binom{n}{k} \exp \left[-\zeta k \log \left(\frac{\zeta k}{\binom{k}{2} \frac{\lambda}{n}} \right) + \zeta k \right] \\ &= \sum_{k \leq n/\log n} \exp \left[(1 - \zeta) k \log \left(\frac{n}{k} \right) + O(k) \right] = o(1). \end{aligned}$$

$\mathbb{P}[(iii) \text{ fails}] = o(1)$ and $\mathbb{P}[(iv) \text{ fails}] = o(1)$ are well-known: indeed, the typical value of maximal degree of a $\mathbf{G}(n, \frac{\lambda}{n})$ graph is of order $\log n / \log \log n$ (see e.g. [22, Theorem 3.4]), and the typical value for the minimal size of connected subgraphs containing more than one cycles is of order $\log n$ (the upper-bound follows readily from a union bound). Finally, for (v), since the expected number of k -cycles in a $\mathbf{G}(n, \frac{\lambda}{n})$ graph is bounded by λ^k for any $k \geq 3$, by Markov inequality we get

$$\mathbb{P}[(v) \text{ fails}] \leq \sum_{k=3}^{\infty} \frac{\lambda^k}{n^{\delta_1 k}} = o(1).$$

Altogether, we conclude the lemma. \square

To make another layer of truncation, we introduce the concept of *good set*: on the one hand, good set is abundant as shown in Lemma 3.13; on the other hand, good set will help controlling both the enumeration of embeddings as in Lemma 3.21 and the number of edges in certain subgraphs as in Lemma 3.23.

Definition 3.3. For a graph $\mathcal{H} = (V, \mathcal{E})$, let $d_{\mathcal{H}}$ be the graph distance on \mathcal{H} . We say a set $A \subset V$ is a *good set* in \mathcal{H} , if it satisfies the following two properties:

- For any two vertices $u, v \in A$, $d_{\mathcal{H}}(u, v) > 2C + 2$.
- For any vertex $w \in A$ and any k -cycle \mathcal{C} in \mathcal{H} with $k \leq C$, we have $d_{\mathcal{H}}(w, \mathcal{C}) > C$.

Denote $K = \lfloor n^\beta \rfloor$. The following lemma allows us to restrict our consideration on good sets with size K in later discussions.

Lemma 3.13. *When n is large enough, for any triple $(\pi^*, G, \mathbb{G}) \in \mathcal{G}$ and any subset $B \subset V$ with $|B| \geq \delta n$, there exists $A \subset B$ with $|A| = K$ such that A is a good set in \mathcal{H}_{π^*} . In other words, for any $B \subset V$ with $|B| \geq \delta n$,*

$$\mathbf{1}_{\mathcal{G}} \leq \sum_{A \subset B, |A|=K} \mathbf{1}_{\mathcal{G}} \mathbf{1}_{\{A \text{ is a good set in } \mathcal{H}_{\pi^*}\}}. \quad (3.32)$$

Proof. Let A be a maximal good set that is contained in B . It suffices to show that $|A| \geq K$ (since if $|A| > K$ we can then take a subset of A of cardinality K which will be a good set). Let $B_0 \subset B$ be the collection of vertices $v \in B$ satisfying $d_{\mathcal{H}_{\pi^*}}(v, \mathcal{C}) \leq C$ for some k -cycle \mathcal{C} in \mathcal{H}_{π^*} with $k \leq C$. For $v \in B$, let R_v be the collection of all vertices $u \in B$ such that $d_{\mathcal{H}}(u, v) \leq 2C + 2$. On the event \mathcal{G} , we have that the maximal degree of \mathcal{H}_{π^*} is bounded by $\log n$, and the number of k -cycles in \mathcal{H}_{π^*} is bounded by $n^{\delta_1 k}$. Therefore,

$$|B_0| \leq \sum_{k=1}^C n^{\delta_1 k} k (\log n)^C \text{ and } |R_v| \leq (\log n)^{2C+2} \text{ for } v \in V. \quad (3.33)$$

Note that B must be contained in $\cup_{v \in A} R_v \cup B_0$, since otherwise one can add a vertex from $B \setminus (\cup_{v \in A} R_v \cup B_0)$ to A which yields a larger good set and thus contradicts to the maximality of A . Combined with the assumption $|B| \geq \delta n$, the choices for β, δ_1 and (3.33), this implies that $|A| \geq K$ for large n . \square

3.3.2 Proof of Proposition 3.11 via second moment bound

Denote the edge likelihood ratio function as

$$\ell(x, y) \stackrel{\text{def}}{=} \frac{\mathcal{Q}[(G_e, \mathbb{G}_{\Pi^*(e)}) = (x, y)]}{\mathbb{P}[G_e = x] \mathbb{P}[\mathbb{G}_{\Pi^*(e)} = y]} = \begin{cases} \frac{1-2ps+ps^2}{(1-ps)^2}, & x = y = 0; \\ \frac{1-s}{1-ps}, & x = 1, y = 0 \text{ or } x = 0, y = 1; \\ \frac{1}{p}, & x = y = 1. \end{cases}$$

Further, write

$$P = \frac{(1 - 2ps + ps^2)}{p(1 - s)^2}, \quad Q = \frac{(1 - s)(1 - ps)}{1 - 2ps + ps^2}, \quad R = \frac{1 - 2ps + ps^2}{(1 - ps)^2}. \quad (3.34)$$

Then it is straightforward to check that

$$\frac{\mathcal{Q}[\pi, G, \mathbb{G}]}{\mathbb{P}[G, \mathbb{G}]} = \frac{1}{n!} \frac{\mathcal{Q}[G, \mathbb{G} \mid \pi^* = \pi]}{\mathbb{P}[G, \mathbb{G}]} = \frac{1}{n!} \prod_{e \in E_0} \ell(G_e, \mathbb{G}_{\Pi(e)}) = \frac{P^{|\mathcal{E}_\pi|} Q^{|E|+|\mathbb{E}|} R^{\binom{n}{2}}}{n!}, \quad (3.35)$$

where \mathcal{E}_π is the set of edges in the π -intersection graph \mathcal{H}_π .

For any $A \subset V$ with $|A| = K$ and any $\sigma \in B(A, \mathbb{V})$, define

$$\begin{aligned} f(G, \mathbb{G}, A, \sigma) &= \sum_{\pi \in B(V, \mathbb{V}, A, \sigma)} P^{|\mathcal{E}_\pi|} Q^{|E|+|\mathbb{E}|} R^{\binom{n}{2}} \mathbf{1}_{\mathcal{G}} \mathbf{1}_{\{A \text{ is good in } \mathcal{H}_\pi\}} \\ &= P^{|\mathcal{E}_\sigma|_A} Q^{|E_A|+|\mathbb{E}_A|} R^{\binom{K}{2}} \sum_{\pi \in B(V, \mathbb{V}, A, \sigma)} P^{|\mathcal{E}_\pi|_A} Q^{|E^A|+|\mathbb{E}^A|} R^{\binom{n}{2} - \binom{K}{2}} \mathbf{1}_{\mathcal{G}} \mathbf{1}_{\{A \text{ is good in } \mathcal{H}_\pi\}}, \end{aligned} \quad (3.36)$$

where $(\mathcal{E}_\sigma)_A$ is the edge set of $(\mathcal{H}_\pi)_A$ for any $\pi \in B(V, \mathbb{V}, A, \sigma)$ (note that this is well-defined). In addition, we set

$$g(G, \mathbb{G}, A) = \max_{\sigma \in B(A, \mathbb{V})} f(G, \mathbb{G}, A, \sigma). \quad (3.37)$$

Proposition 3.14. *For any $\tilde{\pi} \in B(V, \mathbb{V})$, we have*

$$\sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \frac{\mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}}}{\mathbb{P}[G, \mathbb{G}]} \leq \frac{1}{n!} \sum_{A \subset V, |A|=K} f(G, \mathbb{G}, A, \tilde{\pi}|_A). \quad (3.38)$$

Therefore, the left hand side of (3.24) is bounded by

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} \frac{1}{n!} \sum_{A \subset V, |A|=K} g(G, \mathbb{G}, A), \quad (3.39)$$

Proof. For any $\phi \in S_V$, denote $F(\phi)$ for the set of vertices in V fixed by ϕ . Fix some $\tilde{\pi} \in B(V, \mathbb{V})$ and for any $\pi \in B(V, \mathbb{V})$ with $\text{overlap}(\pi, \tilde{\pi}) \geq \delta n$, apply Lemma 3.13 to the set $B = F(\pi^{-1} \circ \tilde{\pi})$ and then sum over all such π , we get that the left hand side of (3.38)

is upper-bounded by

$$\begin{aligned}
& \sum_{\pi: \text{overlap}(\pi, \tilde{\pi}) \geq \delta n} \sum_{A \subset F(\pi^{-1} \circ \tilde{\pi}), |A|=K} \frac{\mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} \mathbf{1}_{\{A \text{ is good in } \mathcal{H}_\pi\}}}{P[G, \mathbb{G}]} \\
& \leq \sum_{A \subset V, |A|=K} \sum_{\pi \in B(V, \mathbb{V}, A, \tilde{\pi}|_A)} \frac{\mathcal{Q}[\pi, G, \mathbb{G}] \mathbf{1}_{\mathcal{G}} \mathbf{1}_{\{A \text{ is good in } \mathcal{H}_\pi\}}}{P[G, \mathbb{G}]} \\
& = \frac{1}{n!} \sum_{A \subset V, |A|=K} f(G, \mathbb{G}, A, \tilde{\pi}|_A),
\end{aligned}$$

where the last equality follows from (3.35) and the definition of $f(G, \mathbb{G}, A, \sigma)$ in (3.36). This proves (3.38), and (3.39) follows immediately from the definition of $g(G, \mathbb{G}, A)$ in (3.37). \square

The main technical input for deriving Proposition 3.11 is incorporated in the following proposition, whose proof is postponed to the appendix.

Proposition 3.15. *For any $A \subset V$ with $|A| = K$ and any $\sigma \in B(A, \mathbb{V})$,*

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} f(G, \mathbb{G}, A, \sigma)^2 \leq ((n - K)!)^2 \exp(\zeta K \log K + \zeta(\alpha - 1)K \log n + o(K \log n)).$$

Proof of Proposition 3.11 assuming Proposition 3.15. By Proposition 3.14, it suffice to show

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} \sum_{A \subset V, |A|=K} g(G, \mathbb{G}, A) = o(n!). \quad (3.40)$$

Applying Cauchy-Schwarz inequality to $\binom{n}{K}$ real numbers $g(G, \mathbb{G}, A)$ (for $A \subset V$ with $|A| = K$), we can upper-bound the left hand side of (3.40) by

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} \left(\binom{n}{K} \sum_{A \subset V, |A|=K} g(G, \mathbb{G}, A)^2 \right)^{1/2}.$$

Use the simple fact that $\max_{x \in S} x^2 \leq \sum_{x \in S} x^2$, the expression above is bounded by

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} \left(\binom{n}{K} \sum_{A \subset V, |A|=K} \sum_{\sigma \in B(A, \mathbb{V})} f(G, \mathbb{G}, A, \sigma)^2 \right)^{1/2}.$$

By Jensen's inequality, we get that the left hide side of (3.40) is further bounded by

$$\left(\binom{n}{K} \sum_{A \subset V, |A|=K} \sum_{\sigma \in B(A, \mathbb{V})} \mathbb{E}_{(G, \mathbb{G}) \sim \mathbb{P}} f(G, \mathbb{G}, A, \sigma)^2 \right)^{1/2}.$$

Combined with Proposition 3.15, this is no more than

$$\begin{aligned} & \left[\binom{n}{K}^3 K!((n-K)!)^2 \exp(\zeta K \log K + \zeta(\alpha-1)K \log n + o(K \log n)) \right]^{1/2} \\ &= n! \exp\left(\frac{1 + \zeta(\alpha-1) - \beta(2-\zeta)}{2} K \log n + o(K \log n)\right), \end{aligned} \quad (3.41)$$

which is $o(n!)$ by the choice of β as in (3.29). This proves (3.40) and thus completes the proof of Proposition 3.11. \square

3.4 Complimentary proofs

3.4.1 Proof of Proposition 3.6

Proof. Fix $\theta > 0$. Let $\{(I_k, J_k, \mathbf{J}_k)\}_{k=1}^\infty$ be i.i.d. triples of independent Bernoulli variables with $\mathbb{E}I_k = p$ and $\mathbb{E}J_k = \mathbb{E}\mathbf{J}_k = s$. For each integer $m \geq 1$, let a_m, b_m, c_m be the value of

$$\mathbb{E} \exp\left(\theta \sum_{i=0}^{m-1} G_i \mathbf{G}_{i+1}\right)$$

where $(G_k, \mathbf{G}_k) = (I_k J_k, I_k \mathbf{J}_k)$ for $1 \leq k \leq m-1$ and $(G_0, \mathbf{G}_m) = (0, 0), (1, 1), (0, 1)$ respectively.

A straightforward recurrence argument yields that for any $m \geq 1$,

$$\begin{cases} a_{m+1} = psc_m + (1-ps)a_m, \\ c_{m+1} = pse^\theta[sc_m + (1-s)a_m] + [ps(1-s)c_m + (1-2ps+ps^2)a_m], \end{cases}$$

and

$$\begin{cases} b_{m+1} = pse^\theta[sb_m + (1-s)c_m] + [ps(1-s)b_m + (1-2ps+ps^2)c_m], \\ c_{m+1} = psb_m + (1-ps)c_m. \end{cases}$$

As a result, a_m, b_m, c_m can be written as linear combinations of μ_1^m and μ_2^m where $\mu_1 > 1 > \mu_2 > 0$ are the two roots of the characteristic polynomial

$$x^2 - (1 + ps^2\nu)x + (ps^2 - p^2s^2)\nu, \quad \text{where } \nu \stackrel{\text{def}}{=} e^\theta - 1.$$

For any θ with $1 \ll e^\theta \ll n$, asymptotically it holds

$$\begin{cases} \mu_1 = 1 + p^2 s^2 \nu + p^3 s^4 \nu^2 + O(p^4 s^4 \nu^2) = 1 + e^\theta / n^{1+\alpha+o(1)}, \\ \mu_2 = ps^2 \nu - p^2 s^2 \nu - p^3 s^4 \nu^2 + O(p^4 s^4 \nu^2) = (\lambda + o(1))e^\theta / n. \end{cases} \quad (3.42)$$

Indeed, $1 \ll e^\theta \ll n$ implies $ps^2 \nu \ll 1 \ll \nu$, and as a result we have

$$\begin{aligned} \mu_1 &= \frac{1 + ps^2 \nu + \sqrt{(1 + ps^2 \nu)^2 - 4(ps^2 - p^2 s^2) \nu}}{2} = \frac{1 + ps^2 \nu + (1 - ps^2 \nu) \sqrt{1 + \frac{4p^2 s^2 \nu}{(1 - ps^2 \nu)^2}}}{2} \\ &= \frac{1 + ps^2 \nu + (1 - ps^2 \nu) \left(1 + \frac{2ps^2 \nu}{(1 - ps^2 \nu)^2} + O(p^4 s^4 \nu^2)\right)}{2} = 1 + p^2 s^2 \nu + p^3 s^4 \nu^2 + O(p^4 s^4 \nu^2). \end{aligned}$$

This yields the first equity in (3.42) and the second follows from $\mu_2 = 1 + ps^2 \nu - \mu_1$.

Note that for any orbit $O_k \in (\mathcal{O}_\pi)_A$ with length k , $\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot|\pi^*]} \exp(\theta |\mathcal{E}_{O_k}|)$ is a linear combination of a_k, b_k and c_k . More precisely, when O is a k -cycle,

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot|\pi^*]} \exp(\theta |\mathcal{E}_{O_k}|) = (1 - 2ps + ps^2)a_k + ps^2 b_k + 2ps(1 - s)c_k,$$

and when O is a k -chain,

$$\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot|\pi^*]} \exp(\theta |\mathcal{E}_{O_k}|) = (1 - ps)^2 a_k + p^2 s^2 b_k + 2ps(1 - ps)c_k.$$

Hence in both cases, $\mathbb{E}_{(G, \mathbb{G}) \sim \mathcal{Q}[\cdot|\pi^*]} \exp(\theta |\mathcal{E}_{O_k}|)$ has the form $c_1 \mu_1^k + c_2 \mu_2^k$ for some constants c_1, c_2 determined by initial conditions.

In the cycle case, it can be shown that $c_1 = c_2 = 1$, and this proves (3.7) (see also Remark 3.16). In the chain case, we can compute the initial values of $k = 1, 2$ explicitly.

Then we obtain

$$\begin{cases} c_1 \mu_1 + c_2 \mu_2 = 1 + p^2 s^2 \nu \triangleq A_1, \\ c_1 \mu_1^2 + c_2 \mu_2^2 = 1 + 2p^2 s^2 \nu + p^3 s^4 \nu^2 \triangleq A_2. \end{cases}$$

Solving the linear system yields that

$$c_1 = \frac{A_2 - \mu_2 A_1}{\mu_1(\mu_1 - \mu_2)}, \quad c_2 = \frac{\mu_1 A_1 - A_2}{\mu_2(\mu_1 - \mu_2)}. \quad (3.43)$$

Plugging (3.42) into (3.43) then gives

$$\begin{aligned} c_1 &= \frac{1 + 2p^2s^2\nu + p^3s^4\nu^2 - (1 + p^2s^2\nu)((ps^2 - p^2s^2)\nu - p^3s^4\nu^2 + O(p^4s^4\nu^2))}{(1 + p^2s^2\nu + p^3s^4\nu^2 + O(p^4s^4\nu^2))(1 - ps^2\nu + 2p^2s^2\nu + 2p^3s^4\nu^2 + O(p^4s^4\nu^2))} \\ &\leq \frac{1 - ps^2\nu + 3p^2s^2\nu + p^3s^4\nu^2 + O(p^4s^4\nu^2)}{1 - ps^2\nu + 3p^2s^2\nu + 3p^3s^4\nu^2 + O(p^4s^4\nu^2)} \leq 1, \end{aligned}$$

and

$$\begin{aligned} c_2 &= \frac{(1 + p^2s^2\nu)(1 + p^2s^2\nu + p^3s^4\nu^2 + O(p^4s^4\nu^2)) - (1 + 2p^2s^2\nu + p^3s^4\nu^2 + O(p^4s^4\nu^2))}{(1 + o(1))ps^2\nu} \\ &= \frac{O(p^4s^4\nu^2)}{(1 + o(1))ps^2\nu} = O(p^3s^2e^\theta) = e^\theta n^{-1-2\alpha+o(1)}. \end{aligned}$$

This completes the proof. \square

Remark 3.16. Another way to see that why the coefficients must be 1 in the cycle case is the following: as illustrated in [44, Appendix A], one can view $\mathbb{E}_{(G, \mathcal{G}) \sim \mathcal{Q}[\cdot|\pi^*]} \exp(\theta|\mathcal{E}_\pi(O_k)|)$ as the trace of the k -th power of certain integral operator \mathcal{L} on the space of real functions on $\{0, 1\}$. The spectrum of \mathcal{L} is precisely $\{\mu_1, \mu_2\}$, so the trace of \mathcal{L}^k equals to $\mu_1^k + \mu_2^k$.

3.4.2 Proof of Lemma 3.9

Proof. For exposition convenience, we prove a slightly stronger statment where we allow x_0, \dots, x_{N+1} to take real values instead of just integer values (but otherwise satisfy the same set of inequalities as specified in the definition of Σ_T and Δ_{n_1, \dots, n_N}). A straightforward algebraic manipulation yields that

$$x_0 + \sum_{k=1}^{N+1} \alpha_k x_k = (1 - \alpha_{N+1})x_0 + \alpha_{N+1} \sum_{k=0}^{N+1} x_k - \sum_{m=1}^N (\alpha_{m+1} - \alpha_m) \sum_{k=1}^m x_k.$$

As a result, the minimum of $x_0 + \sum_{k=1}^{N+1} \alpha_k x_k$ for $(x_0, \dots, x_{N+1}) \in \Sigma_T \cap \Delta_{n_1, \dots, n_N}$ is achieved at $x_0 = 0, x_k = (\varrho + \eta)kn_k, 1 \leq k \leq N$ and $x_{N+1} = [(\alpha - \eta)T - (\varrho + \eta) \sum_{k=1}^N kn_k] \vee 0$.

Hence, $M(T, n_1, \dots, n_N)$ is no less than

$$\begin{aligned}
& [\alpha_{N+1}(\varrho - \eta) - 1]T - \sum_{k=1}^N [\alpha_{N+1}(\varrho + \eta)k - (k-1)(\varrho + \eta) - 1]n_k \\
& \geq [\alpha_{N+1}(\varrho - \eta) - 1]T - \sum_{k: 1 \leq k \leq N, (1-k(1-\alpha_{N+1}))(\varrho + \eta) > 1} [(1-k(1-\alpha_{N+1}))(\varrho + \eta) - 1]n_k \\
& \geq [\alpha_{N+1}(\varrho - \eta) - 1]T - [\alpha_{N+1}(\varrho + \eta) - 1]n_1 \\
& \quad - \max_{2 \leq k \leq \frac{\varrho + \eta - 1}{(1-\alpha_{N+1})(\varrho + \eta)}} \frac{[(1-k(1-\alpha_{N+1}))(\varrho + \eta) - 1]}{k} \times \sum_{t=2}^N tn_t. \tag{3.44}
\end{aligned}$$

If $\frac{\varrho + \eta - 1}{(1-\alpha_{N+1})(\varrho + \eta)} < 2$, then the right hand side of (3.44) is lower-bounded by

$$[\alpha_{N+1}(\varrho - \eta) - 1]T - [\alpha_{N+1}(\varrho + \eta) - 1]n_1 \geq [\alpha_{N+1}(\varrho - \eta) - 1 - c_\lambda^{-1}(\varrho + \eta)\delta]T. \tag{3.45}$$

If $\frac{\varrho + \eta - 1}{(1-\alpha_{N+1})(\varrho + \eta)} \geq 2$, then the right hand side of (3.44) is lower-bounded by

$$\begin{aligned}
& [\alpha_{N+1}(\varrho - \eta) - 1]T - [\alpha_{N+1}(\varrho + \eta) - 1]n_1 - \left[\frac{\varrho + \eta - 1}{2} - (1 - \alpha_{N+1})(\varrho + \eta) \right] T \\
& = \left[\frac{\varrho - \eta - 1}{2} - \left(2\alpha_{N+1} - \frac{1}{2} \right) \eta \right] T - [\alpha_{N+1}(\varrho + \xi) - 1]n_1 \\
& \geq \left[\frac{\varrho - 4\eta - 1}{2} - c_\lambda^{-1}(\varrho + \eta)\delta \right] T. \tag{3.46}
\end{aligned}$$

From (3.45) and (3.46), we may take positive constants δ, δ_0 small enough so that

$$0 < \delta_0 < [\alpha_{N+1}(\varrho - \eta) - 1 - c_\lambda^{-1}(\varrho + \eta)\delta] \wedge \left[\frac{\varrho - 4\eta - 1}{2} - c_\lambda^{-1}(\varrho + \eta)\delta \right],$$

where the right hand side above is positive for small δ due to the choice of η in Definition 3.1 and N in (3.9). This concludes the lemma. \square

3.4.3 Proof of Proposition 3.15

In this section we give the proof of Proposition 3.15. Our method for controlling the truncated second moment is essentially the same as approaches in [12, Section 3], and most steps here are extracted from [12] with only notational changes. However, a straightforward

application of arguments given in [12] is not enough to yield the desired bound, and for this reason we have introduced additional truncations and have considered extra combinatorial structures. Therefore, we provide a self-contained proof here for completeness, despite the fact that the proof enjoys a substantial overlap with that in [12].

Recall the definition of H_A and H^A for a graph $H = (V, E)$ and a subset A of V in Section 3.1.3. Throughout this section, we fix $A \subset V, |A| = K$ and $\sigma \in B(A, \mathbf{V})$. For any (π^*, G, \mathbf{G}) with $\pi^* \in B(V, \mathbf{V}, A, \sigma)$, define

$$\mathcal{G}_A^1 = \{ |(\mathcal{E}_\sigma)_A| \leq \zeta K \}, \quad \mathcal{G}_A^2 = \{ (\mathcal{H}_{\pi^*})^A \text{ is admissible} \}, \quad \mathcal{G}_A^3 = \{ A \text{ is a good set in } (H_{\pi^*})^A \}.$$

Then it is clear that $\mathcal{G} \subset \mathcal{G}_A^1 \cap \mathcal{G}_A^2, \{A \text{ is a good set in } \mathcal{H}_{\pi^*}\} \subset \mathcal{G}_A^3$ and \mathcal{G}_A^1 is independent with $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$ under \mathbf{P} . As a result, the second moment of $f(G, \mathbf{G}, A, \sigma)$ under \mathbf{P} is bounded by

$$\begin{aligned} \mathbb{E}_{(G, \mathbf{G}) \sim \mathbf{P}} f(G, \mathbf{G}, A, \sigma)^2 &\leq \mathbb{E}_{(G_A, \mathbf{G}_A) \sim \mathbf{P}} \left(P^{|(\mathcal{E}_\sigma)_A|} Q^{|E_A| + |\mathbf{E}_A|} R^{\binom{K}{2}} \mathbf{1}_{\mathcal{G}_A^1} \right)^2 \\ &\times \mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathbf{P}} \left(\sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} P^{|(\mathcal{E}_\pi)^A|} Q^{|E^A| + |\mathbf{E}^A|} R^{\binom{n}{2} - \binom{K}{2}} \mathbf{1}_{\mathcal{G}_A^2 \cap \mathcal{G}_A^3} \right)^2. \end{aligned} \quad (3.47)$$

Denote the first and second term in the expression by (I) and (II), respectively. Proposition 3.15 follows immediately once we can upper-bound (I) and (II) as in the next two propositions.

Proposition 3.17. (I) $\leq \exp(\zeta K \log K + \zeta(\alpha - 1)K \log n + o(K \log n))$.

Proposition 3.18. (II) $\leq ((n - K)!)^2 \exp(o(K \log n))$.

We first give the proof of Proposition 3.17, which is a standard computation for the truncated exponential moments for binomial variables.

Proof of Proposition 3.17. Recall the definition of P, Q, R in (3.34). Since $Q \leq 1$ and $R^{2\binom{K}{2}} = \exp(o(K))$, we just need to show that

$$\mathbb{E}_{(G_A, \mathbf{G}_A) \sim \mathbf{P}} P^{2|(\mathcal{E}_\sigma)_A|} \mathbf{1}_{\mathcal{G}_A^1} \leq \exp(\zeta K \log K + \zeta(\alpha - 1)K \log n + o(K \log n)). \quad (3.48)$$

Since $|(\mathcal{E}_\pi)_A| \sim \mathbf{B}(\binom{K}{2}, (ps)^2)$ (that is, a binomial variable with parameters $\binom{K}{2}$ and $(ps)^2$) for $(G_A, \mathbf{G}_A) \sim \mathbf{P}$, the left hand side of (3.48) can be expressed as

$$\begin{aligned} & \sum_{t=0}^{\lfloor \zeta K \rfloor} \binom{\binom{K}{2}}{t} P^{2t} p^{2t} s^{2t} (1 - p^2 s^2)^{\binom{K}{2} - t} \leq \exp(o(K)) \times \sum_{t=1}^{\lfloor \zeta K \rfloor} \binom{\binom{K}{2}}{t} s^{2t} \\ &= \exp(o(K)) \times \binom{\binom{K}{2}}{\lfloor \zeta K \rfloor} s^{2\lfloor \zeta K \rfloor} \times \left(1 + \sum_{t=1}^{\lfloor \zeta K \rfloor} \prod_{i=0}^{t-1} \frac{\lfloor \zeta K \rfloor - i}{(\binom{K}{2} - \lfloor \zeta K \rfloor + i) s^2} \right). \end{aligned} \quad (3.49)$$

Since $\beta > 1 - \alpha$ by (3.29), $Ks^2 \gg 1$ and the last term in (3.49) is $1 + o(1)$, we see that the left hand side of (3.48) equals to $\exp(\zeta K \log K + \zeta(\alpha - 1)K \log n + o(K \log n))$, as desired. \square

The rest of this section is devoted to the proof of Proposition 3.18. Denote \mathbf{A} for the image of A under σ . Recall $E_0^A \subset E_0$ is the set of edges in E_0 not within A , and similarly for \mathbf{E}_0^A . We now define several probability measures $\mathcal{Q}_{A, \mathbf{A}, \sigma}, \mathcal{Q}'_{A, \mathbf{A}, \sigma}, \mathbf{P}_{A, \mathbf{A}}, \mathbf{Q}_{A, \mathbf{A}, \sigma}, \mathbf{Q}'_{A, \mathbf{A}, \sigma}$, which will play important roles in our proof.

Definition 3.4. Let $\mathcal{Q}_{A, \mathbf{A}, \sigma}$ be the probability measure on the space

$$\Omega_{A, \mathbf{A}} = \{(\pi^*, G^A, \mathbf{G}^A) : \pi^* \in B(V, \mathbf{V}), G^A, \mathbf{G}^A \text{ are subgraphs of } (V, E_0^A), (V, \mathbf{E}_0^A)\}$$

defined as follow: the marginal distribution of π^* under $\mathcal{Q}_{A, \mathbf{A}, \sigma}$ is uniform on $B(V, \mathbf{V}, A, \sigma)$, and conditioned on π^* , (G^A, \mathbf{G}^A) is obtained by deleting edges within A and \mathbf{A} from a pair of correlated Erdős-Rényi graphs (G, \mathbf{G}) sampled according to $\mathcal{Q}[\cdot \mid \pi^*]$. Let $\mathcal{Q}'_{A, \mathbf{A}, \sigma}$ be the conditional law of $\mathcal{Q}_{A, \mathbf{A}, \sigma}$ under the event $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$.

Further, define $\mathbf{P}_{A, \mathbf{A}}$ as the law of a pair of independent Erdős-Rényi graphs with edge density ps on (V, E_0^A) and (V, \mathbf{E}_0^A) , and $\mathbf{Q}_{A, \mathbf{A}, \sigma}, \mathbf{Q}'_{A, \mathbf{A}, \sigma}$ as the marginal law for the last two coordinates of $\mathcal{Q}_{A, \mathbf{A}, \sigma}, \mathcal{Q}'_{A, \mathbf{A}, \sigma}$, respectively.

Again, it is straightforward to check that

$$\frac{\mathcal{Q}[\pi, G^A, \mathbf{G}^A]}{\mathbf{P}[G^A, \mathbf{G}^A]} = \frac{1}{(n - K)!} \prod_{e \in E_0^A} \ell(G_e, \mathbf{G}_{\Pi(e)}) = \frac{P^{|\mathcal{E}_\pi|^A} Q^{|E^A| + |\mathbf{E}^A|} R^{\binom{n}{2} - \binom{K}{2}}}{(n - K)!}.$$

Combined with the fact that for any triple $(\pi^*, G^A, \mathbf{G}^A) \in \Omega_{A,A}$ we have $\mathcal{Q}[\pi^*, G^A, \mathbf{G}^A] \mathbf{1}_{\mathcal{G}_A^2 \cap \mathcal{G}_A^3} \leq \mathcal{Q}'_{A,A,\sigma}[\pi^*, G^A, \mathbf{G}^A]$, it yields that the term (II) is bounded by

$$(n-K)! \sum_{\pi|_A=\sigma} \frac{\mathcal{Q}_{A,A,\sigma}[\pi, G^A, \mathbf{G}^A] \mathbf{1}_{\mathcal{G}_A^2 \cap \mathcal{G}_A^3}}{\mathbf{P}_{A,A}[G^A, \mathbf{G}^A]} \leq (n-K)! \frac{\mathcal{Q}'_{A,A,\sigma}[G^A, \mathbf{G}^A]}{\mathbf{P}_{A,A}[G^A, \mathbf{G}^A]}.$$

It suffice to bound the second moment of the conditional likelihood ratio $L'(G^A, \mathbf{G}^A) \stackrel{\text{def}}{=} \frac{\mathcal{Q}'_{A,A,\sigma}[G^A, \mathbf{G}^A]}{\mathbf{P}_{A,A}[G^A, \mathbf{G}^A]}$ under \mathbf{P} , or equivalently, under $\mathbf{P}_{A,A}$. Note that the conditional law of L' given π^* is invariant of the realization of π^* . So in what follows, we may assume π^* as certain fixed element in $B(V, \mathbf{V}, A, \sigma)$.

Recall the definition of the set of edge orbits \mathcal{O}_π in E_0 induced by $\phi = \pi^{-1} \circ \pi^*$. For any $\pi \in B(V, \mathbf{V}, A, \sigma)$, let $\mathcal{O}_\pi^A \subset \mathcal{O}_\pi$ be the edge orbits that are entirely contained in E_0^A and let $\mathcal{J}_\pi \subset \mathcal{O}_\pi^A$ be the set of edge orbits in \mathcal{O}_π^A that are *entirely* contained in $(\mathcal{H}_{\pi^*})^A$ (i.e., the π^* -intersecting graph of G^A and \mathbf{G}^A). Note that while \mathcal{O}_π^A is deterministic whenever π is fixed, \mathcal{J}_π is random depending on the realization of (G^A, \mathbf{G}^A) . Let $H(\mathcal{J}_\pi)$ be the subgraph of $(\mathcal{H}_{\pi^*})^A$ with vertices and edges from orbits in \mathcal{J}_π . With slight abuse of notation, we denote $|\mathcal{J}_\pi|$ for the total number of *edges* in orbits $O \in \mathcal{J}_\pi$. Inspired by [12, Lemma 3.2], we have the following proposition, which connects the second moment of L' with the exponential moment of \mathcal{J}_π .

Proposition 3.19. *The second moment of $L'(G^A, \mathbf{G}^A)$ under $\mathbf{P}_{A,A}$ is bounded by*

$$\frac{1}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_2^A \cap \mathcal{G}_3^A]} \mathbb{E}_{(\pi^*, G^A, \mathbf{G}^A) \sim \mathcal{Q}'_{A,A,\sigma}} \frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \frac{p^{-|\mathcal{J}_\pi|}}{\mathcal{Q}[\mathcal{G}_2^A \cap \mathcal{G}_3^A \mid \pi^*, \mathcal{J}_\pi]}. \quad (3.50)$$

Proof. It is clear that

$$\begin{aligned} \mathcal{Q}'_{A,A,\sigma}[G^A, \mathbf{G}^A] &= \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \mathcal{Q}'_{A,A,\sigma}[\pi, G^A, \mathbf{G}^A] \\ &\leq \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \frac{\mathcal{Q}_{A,A,\sigma}[\pi, G^A, \mathbf{G}^A]}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_2^A \cap \mathcal{G}_3^A]} = \frac{\mathcal{Q}_{A,A,\sigma}[G^A, \mathbf{G}^A]}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_2^A \cap \mathcal{G}_3^A]}. \end{aligned}$$

Thus, $L'(G^A, \mathbf{G}^A)$ is bounded by $\frac{L(G^A, \mathbf{G}^A)}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3]}$, where $L(G^A, \mathbf{G}^A) \stackrel{\text{def}}{=} \frac{\mathcal{Q}_{A,A,\sigma}[G^A, \mathbf{G}^A]}{\mathcal{P}_{A,A}[G^A, \mathbf{G}^A]}$ is the unconditional likelihood ratio and can be written explicitly as

$$\begin{aligned} L(G^A, \mathbf{G}^A) &= \frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \prod_{e \in E_0^A} \ell(G_e, \mathbf{G}_{\Pi(e)}) \\ &= \frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \prod_{O \in \mathcal{O}_\pi^A} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}). \end{aligned}$$

Hence, the second moment of L' under $\mathcal{P}_{A,A}$, or equivalently, the first moment of L' under $\mathcal{Q}'_{A,A,\sigma}$ is bounded by

$$\frac{1}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3]} \mathbb{E}_{\pi^* \sim \mathcal{Q}'} \frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} \mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*]} \prod_{O \in \mathcal{O}_\pi^A} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}). \quad (3.51)$$

A crucial fact is that for each $\pi \in B(V, \mathbf{V}, A, \sigma)$ and each orbit $O \in \mathcal{O}_\pi^A$, it holds

$$\mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}_{A,A,\sigma}[\cdot | \pi^*]} \left[\prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \mid O \notin \mathcal{J}_\pi \right] \leq 1. \quad (3.52)$$

In other words, for those cycles not entirely contained in $(\mathcal{H}_{\pi^*})^A$, their contribution to the likelihood ratio is negligible. The proof of (3.52) can be found in [12, Lemma 3.1] via an explicit computation. Intuitively, (3.52) tells us that the main contribution of (3.51) comes from orbits in \mathcal{J}_π and this motivates us to take conditional expectation with respect to the random set \mathcal{J}_π . For any two fixed $\pi^*, \pi \in B(V, \mathbf{V}, A, \sigma)$, by averaging over conditional expectation given \mathcal{J}_π ,

$$\begin{aligned} &\mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}_{A,A,\sigma}[\cdot | \pi^*]} \prod_{O \in \mathcal{O}_\pi^A} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \\ &= \mathbb{E}_{\mathcal{J}_\pi \sim \mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*]} \left[p^{-|\mathcal{J}_\pi|} \mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*]} \left[\prod_{O \in \mathcal{O}_\pi^A \setminus \mathcal{J}_\pi} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \mid \mathcal{J}_\pi \right] \right] \quad (3.53) \end{aligned}$$

For any realization J of \mathcal{J}_π , conditioned on $\mathcal{J}_\pi = J$, we may upper-bound the expectation in the inner layer of (3.53) by

$$\frac{1}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 \mid \pi^*, \mathcal{J}_\pi = J]} \mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}_{A,A,\sigma}[\cdot | \pi^*, \mathcal{J}_\pi = J]} \left[\prod_{O \in \mathcal{O}_\pi^A \setminus J} \prod_{e \in O} \ell(G_e, \mathbf{G}_{\Pi(e)}) \right], \quad (3.54)$$

where the term $\frac{1}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 | \pi^*, \mathcal{J}_\pi = J]}$ emerges because we move from $\mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*, \mathcal{J}_\pi = J]$ to $\mathcal{Q}_{A,A,\sigma}[\cdot | \pi^*, \mathcal{J}_\pi = J]$ in the expectation. Furthermore, note that under the law $\mathcal{Q}_{A,A,\sigma}[\cdot | \pi^*, \mathcal{J}_\pi = J]$, for distinct orbits $O \in \mathcal{O}_\pi^A \setminus J$, the families of random variable $\{(G_e, \mathbf{G}_{\Pi(e)}) : e \in O\}$ are mutually independent with law $\mathcal{Q}_{A,A,\sigma}[\cdot | O \notin \mathcal{J}_\pi]$. Thus the last expectation in (3.54) is no more than 1 by (3.52). Combined this with (3.53) and (3.54) gives that the left hand side of (3.53) is bounded by

$$\mathbb{E}_{\mathcal{J}_\pi \sim \mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*]} \frac{p^{-|\mathcal{J}_\pi|}}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 | \pi^*, \mathcal{J}_\pi]} = \mathbb{E}_{(G^A, \mathbf{G}^A) \sim \mathcal{Q}'_{A,A,\sigma}[\cdot | \pi^*]} \frac{p^{-|\mathcal{J}_\pi|}}{\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 | \pi^*, \mathcal{J}_\pi]} . \quad (3.55)$$

Plugging (3.55) into (3.51) yields the desired result (3.50). \square

The following lemma bounds the probabilities appearing in (3.50) from below.

Lemma 3.20. *There exists a constant $c_0 = c_0(\lambda) > 0$ such that for large enough n ,*

$$\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3] \geq c_0^K . \quad (3.56)$$

Further, for any $\pi^*, \pi \in B(V, \mathbb{V}, A, \sigma)$ and any realization J of \mathcal{J}_π satisfying $\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 | H(J) \subset (\mathcal{H}_{\pi^*})^A] > 0$, it holds

$$\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 | \pi^*, \mathcal{J}_\pi = J] \geq c_0^{K+|J|} . \quad (3.57)$$

Proof. Clearly, both \mathcal{G}_A^2 and \mathcal{G}_A^3 are decreasing events measurable with respect to \mathcal{H}_{π^*} . It can be shown in the same way as Lemma 3.12 that $\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2] = 1 - o(1)$, and

$$\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^3] \geq \mathcal{Q}_{A,A,\sigma}[\{\text{There is no edge between } A \text{ and } V \setminus A \text{ in } (\mathcal{H}_{\pi^*})^A\}] \geq c_1^K$$

for some constant $c_1 = c_1(\lambda) > 0$. Applying FKG inequality to the events \mathcal{G}_A^2 and \mathcal{G}_A^3 yields that $\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3] \geq (1 - o(1))c_1^K$.

Further, $\{\mathcal{J}_\pi = J\}$ can be written as the intersection of $\{H(J) \subset (\mathcal{H}_{\pi^*})^A\}$ (here \subset means being a subgraph of) and a decreasing event \mathcal{G}_J^4 (which is measurable with respect to edges not in $E(J)$). Let \mathcal{A} be the event that there is no edge incident to A except those edges

in $E(J)$. Since $\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^3 \mid H(J) \subset (\mathcal{H}_{\pi^*})^A] > 0$, we have $\mathcal{A} \cap \{H(J) \subset (\mathcal{H}_{\pi^*})^A\} \subset \mathcal{G}_A^3$. Thus,

$$\begin{aligned} \mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 \mid \pi^*, \mathcal{J}_\pi = J] &\geq \mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{A} \mid \pi^*, \mathcal{J}_\pi = J] \\ &\geq \mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{A} \mid \pi^*, H(J) \subset (H_{\pi^*})^A], \end{aligned}$$

where for the second inequality we have applied FKG inequality to the conditional measure $\mathcal{Q}_{A,A,\sigma}[\cdot \mid \pi^*, H(J) \subset (H_{\pi^*})^A]$ (which is a product measure outside of $E(J)$) and the fact that \mathcal{G}_J^4 , \mathcal{G}_A^2 and \mathcal{A} remain to be decreasing events restricted to the space satisfying $H(J) \subset (H_{\pi^*})^A$ (i.e., in the space for the conditional measure $\mathcal{Q}_{A,A,\sigma}[\cdot \mid \pi^*, H(J) \subset (H_{\pi^*})^A]$). Applying FKG again, we get that

$$\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 \mid \pi^*, \mathcal{J}_\pi = J] \geq \mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \mid \pi^*, H(J) \subset (\mathcal{H}_{\pi^*})^A] \mathcal{Q}_{A,A,\sigma}[\mathcal{A} \mid \pi^*, H(J) \subset (\mathcal{H}_{\pi^*})^A].$$

Also, by a similar argument as in the proof of [12, Lemma 3.3] we can show that

$$\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \mid \pi^*, H(J) \subset (\mathcal{H}_{\pi^*})^A] \geq (1 - o(1))c_2^{|J|}$$

for some constant $c_2 = c_2(\lambda) > 0$ (indeed, this is where we use (iv) and (v) in admissibility).

Combined with the preceding inequality, it yields the lemma by picking $c_0 = c_1 c_2 / 2$. \square

Note that any possible realization J for $\mathcal{J}_\pi \sim \mathcal{Q}'_{A,A,\sigma}$ satisfies $\mathcal{Q}_{A,A,\sigma}[\mathcal{G}_A^2 \cap \mathcal{G}_A^3 \mid H(J) \subset (\mathcal{H}_{\pi^*})^A] > 0$. From Lemma 3.20, we see (3.50) is bounded by $\exp(O(K))$ times

$$\mathbb{E}_{(\pi^*, G^A, \mathcal{G}^A) \sim \mathcal{Q}'_{A,A,\sigma}} \frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} (c_0 p)^{-|\mathcal{J}_\pi|}. \quad (3.58)$$

Hence in order to prove Proposition 3.18 it suffices to show that (3.58) is $\exp(o(K \log n))$. The idea is to bound the term in the expectation deterministically for any $(H_{\pi^*})^A$ satisfying $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$. Our strategy is to enumerate all possible realization J of \mathcal{J}_π and control the number of π such that $\mathcal{J}_\pi = J$, then sum over J . To this end, we need to introduce several notations.

For two finite simple graphs H and \mathcal{H} , a *labeled embedding* of H into \mathcal{H} is an injective map $\iota : H \rightarrow \mathcal{H}$, such that $(\iota(u), \iota(v))$ is an edge of \mathcal{H} when (u, v) is an edge of H . Further, we define an *unlabeled embedding* of H into \mathcal{H} as an isomorphic class of labeled embeddings: two labeled embeddings $\iota_1, \iota_2 : H \rightarrow \mathcal{H}$ is said to be equivalent if and only if there is an automorphism $\phi : H \rightarrow H$, such that $\iota_2 = \iota_1 \circ \phi$. Let $t(H, \mathcal{H})$ be the total number of unlabeled embedding of H into \mathcal{H} . Then the total number of labeled embeddings of H into \mathcal{H} is given by $\text{Aut}(H)t(H, \mathcal{H})$, where $\text{Aut}(H)$ denotes the number of automorphisms of H to itself. For each isomorphic class of finite simple graphs \mathcal{C} , pick a representative element $H_{\mathcal{C}}$ of \mathcal{C} and fix it.

Now let $\mathcal{H} = (H_{\pi^*})^A = (V, (\mathcal{E}_{\pi^*})^A)$ be the π^* -intersection graph of G^A and \mathbf{G}^A . Let \mathfrak{H}_{ℓ} be the collection of connected subgraphs of \mathcal{H} which contain ℓ vertices in $V \setminus A$ and at least one vertex in A , let \mathfrak{C}_{ℓ} (respectively \mathfrak{T}_{ℓ}) be the collection of all such representatives $H_{\mathcal{C}}$ which are connected non-tree subgraphs (respectively trees) with ℓ vertices so that $t(H_{\mathcal{C}}, \mathcal{H}) > 0$. Note that by definition, any two graphs in \mathfrak{C}_{ℓ} or \mathfrak{T}_{ℓ} are non-isomorphic, while two graphs in \mathfrak{H}_{ℓ} might be isomorphic. For each $H \in \mathfrak{H}_{\ell}$, fix a vertex $v_H \in H \cap A$ and denote by $\text{Aut}(H, v_H)$ the number of automorphisms of H to itself fixing v_H . When \mathcal{H} is admissible, we have the following bounds on subgraph counts of $\mathfrak{H}_{\ell}, \mathfrak{C}_{\ell}$ and \mathfrak{T}_{ℓ} .

Lemma 3.21. *Under the condition $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$, we have for any $\ell \geq 2$, the following hold:*

$$\sum_{H \in \mathfrak{H}_{\ell}} \text{Aut}(H, v_H) \leq 2K(2^{\xi} \log n)^{4\ell}, \quad (3.59)$$

$$\sum_{C \in \mathfrak{C}_{\ell}} \text{Aut}(C)t(H, \mathcal{H}) \leq \ell^3(2^{\xi+1}n^{\delta_1})^{\ell}, \quad (3.60)$$

$$\sum_{T \in \mathfrak{T}_{\ell}} \text{Aut}(T)t(T, \mathcal{H}) \leq n(4 \log n)^{2(\ell-1)}. \quad (3.61)$$

Proof. (3.60) and (3.61) are proved in the same manner as in [12, Lemma 3.5]. The proof of (3.59) also largely shares the same method with that in [12, Lemma 3.5]. But since a modification is required, in what follows we provide a proof for (3.59). Note that $\sum_{H \in \mathfrak{H}_{\ell}} \text{Aut}(H, v_H)$ is just the total number for labeled embedding of H into \mathcal{H} fixing v_H .

Each labeled embedding of $H \in \mathfrak{H}_\ell$ into \mathcal{H} can be constructed as follow: (1) choose a labeled spanning tree T rooted in A which intersects with $V \setminus A$ on ℓ vertices; and (2) add edges in \mathcal{H} within the vertex set of T to T and get the final labeled embedding. By \mathcal{G}_A^3 , each vertex $v \in V \setminus A$ is connected with at most one vertex in A , since otherwise there would be two vertices in A having distance 2 which is a contradiction to good set. Thus the size of T in the first step is bounded by 2ℓ . Since each labeled tree with no more than 2ℓ vertices can be encoded by a contour (according to depth-first search) with length no more than 4ℓ starting from A , the number of choices for T is bounded by $2K(\log n)^{4\ell}$ on the event \mathcal{G}_A^2 (recall (iii) in admissibility). When T is fixed, there are no more than $2\xi\ell$ edges in \mathcal{H} within the vertex set of T by (i) in admissibility, so the number of ways for adding edges to T is bounded by $2^{2\xi\ell}$. This gives the proof of (3.59). \square

For any realization J of \mathcal{J}_π , J can be decomposed into connected components, some of them intersect A while others do not. We consider a *configuration* Λ consisting of non-negative integers r, s, t , positive integers l_j, x_j, m_k, y_k for $j \in [s], k \in [t]$ and distinct graphs $H_i \in \bigcup_{\ell \geq 2} \mathfrak{H}_\ell$ for $i \in [r]$, $C_j \in \mathfrak{C}_{l_j}$ for $j \in [s]$ and $T_k \in \mathfrak{T}_{m_k}$ for $k \in [t]$. (There is nothing mathematically deep in *configuration*, and this is mainly for notational convenience.) For each such configuration Λ , define $\Pi(\Lambda)$ to be the set of $\pi \in B(V, \mathbf{V}, A, \sigma)$ such that the components of \mathcal{J}_π which intersect with A are exactly $H_i, i \in [r]$, and the other components consist of x_j copies of C_j for $j \in [s]$ and y_k copies of T_k for $k \in [t]$. The next lemma provides an upper bound on $\Pi(\Lambda)$.

Lemma 3.22. *With aforementioned notations, we have*

$$\begin{aligned}
|\Pi(\Lambda)| &\leq \left(n - K - \sum_{i \in [r]} |H_i \setminus A| - \sum_{j \in [s]} x_j l_j - \sum_{k \in [t]} y_k m_k \right)! \times \prod_{i \in [r]} \text{Aut}(H_i, v_{H_i}) \\
&\times \prod_{j \in [s]} (\text{Aut}(C_j) t(C_j, \mathcal{H}))^{x_j} \times \prod_{k \in [t]} (\text{Aut}(T_k) t(T_k, \mathcal{H}))^{y_k}.
\end{aligned} \tag{3.62}$$

Proof. Let H_Λ be a graph which is the disjoint union of H_i for $i \in [r]$, x_j copies of C_j for $j \in [s]$ and y_k copies of T_k for $k \in [t]$. First we choose an unlabeled embedding $\iota : H_\Lambda \rightarrow \mathcal{H}$

with $\iota|_{H_i} = \text{id}, \forall i \in [r]$ and $\iota(H_\Lambda \setminus \bigcup_{i \in [r]} H_i) \cap A = \emptyset$. The number of such choices is bounded by

$$\prod_{j \in [s]} \binom{t(C_j, \mathcal{H})}{x_j} \prod_{k \in [t]} \binom{t(T_k, \mathcal{H})}{y_k} \leq \prod_{j \in [s]} \frac{t(C_j, \mathcal{H})^{x_j}}{x_j!} \prod_{k \in [t]} \frac{t(T_k, \mathcal{H})^{y_k}}{y_k!}.$$

For fixed ι , any $\pi \in B(V, \mathbb{V}, A, \sigma)$ such that $\mathcal{J}_\pi = \iota(H_\Lambda)$ can be decomposed into two permutations π_1 and π_2 on $A \cup \iota(H_\Lambda)$ and $V \setminus (A \cup \iota(H_\Lambda))$, respectively. The number of choices for π_2 is at most

$$\left(n - K - \sum_{i \in [r]} |H_i \setminus A| - \sum_{j \in [s]} x_j l_j - \sum_{k \in [t]} y_k m_k \right)!.$$

For π_1 , we have the following crucial observation: for any $u \in \iota(H_\Lambda)$, π_1 inhibits to an isomorphism between the components of $\iota(H_\Lambda)$ containing u and $\pi_1(u)$. That is to say, for any $u, v \in \iota(H_\Lambda)$, whenever v is adjacent to u , $\pi_1(v)$ is adjacent to $\pi_1(u)$ and $\pi_1^{-1}(v)$ is adjacent to $\pi_1^{-1}(u)$. This is true because by definition, the edge orbit containing (u, v) is entirely contained in \mathcal{J}_π . With such observation and the condition that $\pi_1|_A = \sigma$, we see that π_1 inhibits to an automorphism of H_i to itself fixing v_{H_i} for each $i \in [r]$. In addition, for any C_j (and similarly for T_k), we will “permute” its x_j copies of C_j such that π_1 maps one copy to its image under the permutation, and within each copy we also have the freedom of choosing an arbitrary automorphism of C_j . Since the choices for the permutations and automorphisms determine π_1 , we conclude from this that the number of choices for π_1 is no more than

$$\prod_{i \in [r]} \text{Aut}(H_i, v_{H_i}) \prod_{j \in [s]} x_j! \text{Aut}(C_j)^{x_j} \prod_{k \in [t]} y_k! \text{Aut}(T_k)^{y_k}.$$

These bounds altogether yield (3.62). \square

We need yet another lemma which bounds the number of edges in $H \in \mathfrak{H}_\ell$.

Lemma 3.23. *There exists some constant $\delta_2 > 0$, such that for any constant $c > 0$ and n large enough, under the condition $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$ it holds*

$$\sup_{H \in \mathfrak{H}_\ell} \frac{(2^\xi \log n)^{4\ell}}{n^\ell (cp)^{|E(H)|}} \leq n^{-\delta_2 \ell}, \forall \ell \geq 1. \quad (3.63)$$

Proof. We remark that the proof of this lemma illustrates the motivation for defining good set.

For any $\ell \geq 1$ and $H \in \mathfrak{H}_\ell$, denote $I = H \setminus A$ and let $J \subset I$ be the set of vertices in I with an edge connecting to A in H . Then $|I| = \ell$ by definition of \mathfrak{H}_ℓ , and any vertex in J is adjacent to exactly one vertex in A as before. Further, we have the observation that for any two vertices $u, v \in J$, the C -neighborhoods of u, v are disjoint (since otherwise there would be two vertices in A with graph distance no more than $2C + 2$, which contradicts with \mathcal{G}_A^3). Our proof of (3.63) is divided into three cases.

- *Case 1:* $1 \leq \ell \leq C$. In this case, H must be a tree, since otherwise there would be a cycle with length no more than C and within distance C from A , which contradicts with \mathcal{G}_A^3 . Further, J must be a singleton by the previous observation. Hence $E(H) \leq \ell$ for any $H \in \mathfrak{H}_\ell$ and so

$$\sup_{H \in \mathfrak{H}_\ell} \frac{(2^\xi \log n)^{4\ell}}{n^\ell (cp)^{|E(H)|}} \leq n^{(\alpha - 1 + o(1))\ell}. \quad (3.64)$$

- *Case 2:* $C < \ell \leq n/\log n$. In this case, the C -neighborhood of each $v \in H$ contains at least C vertices, so $|J| \leq \ell/C$. In addition, the edges within I is bounded by $\zeta\ell$ from (ii) of admissibility, so $|E(H)| \leq (\zeta + C^{-1})\ell$ for any $H \in \mathfrak{H}_\ell$. Thus

$$\sup_{H \in \mathfrak{H}_\ell} \frac{(2^\xi \log n)^{4\ell}}{n^\ell (cp)^{|E(H)|}} \leq n^{(\alpha(\zeta + C^{-1}) - 1 + o(1))\ell}. \quad (3.65)$$

- *Case 3:* $\ell \geq n/\log n$. In this case, the number of edges within H is bounded by $\xi|H| \leq \xi(\ell + K)$ from (i) in admissibility, so

$$\sup_{H \in \mathfrak{H}_\ell} \frac{(2^\xi \log n)^{4\ell}}{n^\ell (cp)^{|E(H)|}} \leq n^{\alpha\xi(K+\ell)-\ell}. \quad (3.66)$$

By the choice of ξ, ζ and C in (3.27), (3.28), (3.30), we may choose a positive constant $\delta_2 < (1 - \alpha(\zeta + C^{-1})) \wedge (1 - \alpha\xi)$. Then (3.63) follows from (3.64), (3.65) and (3.66) (note that in the last case $\ell \gg K$). This completes the proof. \square

We are now ready to state and prove the final proposition in this section.

Proposition 3.24. *Whenever $(\pi^*, G^A, \mathcal{G}^A)$ is sampled from $\mathcal{Q}'_{A, \mathbf{A}, \sigma}$, it holds that*

$$\frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} (c_0 p)^{-|\mathcal{J}_\pi|} = \exp(o(K \log n)). \quad (3.67)$$

Provided with Proposition 3.24, we have that (3.58) is $\exp(o(K \log n))$ and thus Proposition 3.17 follows.

Proof. For $(\pi^*, G^A, \mathcal{G}^A) \sim \mathcal{Q}'_{A, \mathbf{A}, \sigma}$, the π^* -intersection graph \mathcal{H} satisfies $\mathcal{G}_A^2 \cap \mathcal{G}_A^3$. We have

$$\frac{1}{(n-K)!} \sum_{\pi \in B(V, \mathbf{V}, A, \sigma)} (c_0 p)^{-|\mathcal{J}_\pi|} \leq \frac{1}{(n-K)!} \sum_{\Lambda} (c_0 p)^{-|E(H_\Lambda)|} \times |\Pi(\Lambda)|, \quad (3.68)$$

where the sum is taken over all possible configurations Λ . It is clear that each tree $T \in \bigcup \mathfrak{T}_\ell$ has $|T| - 1$ edges and each $C \in \bigcup \mathfrak{C}_\ell$ has no more than $\xi|C|$ edges by \mathcal{G}_A^2 and (i) in admissibility. From this and Lemma 3.22, we can upper bound (3.68) by

$$\begin{aligned} & \frac{1}{(n-K)!} \sum_{r, s, t \geq 0} \sum_{H_1, \dots, H_r \in \bigcup \mathfrak{H}_\ell} \sum_{C_1, \dots, C_s \in \bigcup \mathfrak{C}_\ell} \sum_{T_1, \dots, T_t \in \bigcup \mathfrak{T}_\ell} \sum_{x_1, \dots, x_s > 0} \sum_{y_1, \dots, y_t > 0} \\ & (c_0 p)^{-\sum_{i \in [r]} |E(H_i)| - \xi \sum_{j \in [s]} x_j |C_j| - \sum_{k \in [t]} y_k (|T_k| - 1)} \\ & \times \left(n - K - \sum_{i \in [r]} |H_i \setminus A| + \sum_{j \in [s]} x_j l_j + \sum_{k \in [t]} y_k m_k \right)! \\ & \times \prod_{i \in [r]} \text{Aut}(H_i, v_{H_i}) \times \prod_{j \in [s]} (\text{Aut}(C_j) t(C_j, \mathcal{H}))^{x_j} \times \prod_{k \in [t]} (\text{Aut}(T_k) t(T_k, \mathcal{H}))^{y_k}. \end{aligned}$$

Note that by Stirling's formula, with $K = \lfloor n^\beta \rfloor$ we have $(n - K - t)! / (n - K)! \leq (2e/n)^t$ for any $t \geq 0$, thus for $c_1 = c_0/2e$, the expression above is bounded by

$$\begin{aligned} & \sum_{r, s, t \geq 0} \sum_{H_1, \dots, H_r \in \bigcup \mathfrak{H}_\ell} \sum_{C_1, \dots, C_s \in \bigcup \mathfrak{C}_\ell} \sum_{T_1, \dots, T_t \in \bigcup \mathfrak{T}_\ell} \sum_{x_1, \dots, x_s > 0} \sum_{y_1, \dots, y_t > 0} \\ & \prod_{i \in [r]} n^{-|H_i \setminus A|} (c_1 p)^{|E(H_i)|} \text{Aut}(H_i, v_{H_i}) \prod_{j \in [s]} (c_1 n p)^{-\xi x_j |C_j|} \text{Aut}(C_j)^{x_j} \prod_{k \in [t]} (c_1 n p)^{-y_k (|T_k| - 1)} \text{Aut}(T_k)^{y_k} \\ & = \prod_{H \in \bigcup \mathfrak{H}_\ell} \left(1 + \frac{\text{Aut}(H, v_H)}{n^{|H \setminus A|} (c_1 p)^{|E(H)|}} \right) \prod_{C \in \bigcup \mathfrak{C}_\ell} \sum_{x \geq 0} \left(\frac{\text{Aut}(C)}{(c_1 n p)^{\xi |C|}} \right)^x \prod_{T \in \bigcup \mathfrak{T}_\ell} \sum_{y \geq 0} \left(\frac{p \text{Aut}(T)}{(c_1 n p)^{|T|}} \right)^y. \end{aligned} \quad (3.69)$$

As in the proof of [12, Proposition 3.6], it can be shown by (3.60) and (3.61) that the last two terms in (3.69) are $1+o(1)$, and thus it remains to show the first term is $\exp(o(K \log n))$. Since $\log(1+x) \leq x$ for any $x \geq 0$, the logarithm of the first term in (3.69) is bounded by

$$\begin{aligned} & \sum_{\ell \geq 1} \sum_{H \in \mathfrak{H}_\ell} \frac{\text{Aut}(H, v_H)}{n^\ell (c_1 p)^{|E(H)|}} \leq \sum_{\ell \geq 1} \sup_{H \in \mathfrak{H}_\ell} \frac{1}{n^\ell (c_1 p)^{|E(H)|}} \sum_{H \in \mathfrak{H}_\ell} \text{Aut}(H, v_H) \\ & \stackrel{(3.59)}{\leq} K \sum_{\ell \geq 1} \sup_{H \in \mathfrak{H}_\ell} \frac{(2^\xi \log n)^{4\ell}}{n^\ell (c_1 p)^{|E(H)|}} \stackrel{(3.63)}{\leq} K \sum_{\ell \geq 1} n^{-\delta_2 \ell} = o(K \log n), \end{aligned} \tag{3.70}$$

as desired. \square

References

- [1] *HLT '05: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, USA, 2005. Association for Computational Linguistics.
- [2] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.
- [3] V. Anantharam and J. Salez. The densest subgraph problem in sparse random graphs. *Ann. Appl. Probab.*, 26(1):305–327, 2016.
- [4] R. Arratia and S. Tavaré. The Cycle Structure of Random Permutations. *The Annals of Probability*, 20(3):1567 – 1591, 1992.
- [5] B. Barak, C.-N. Chou, Z. Lei, T. Schramm, and Y. Sheng. (nearly) efficient algorithms for the graph matching problem on correlated random graphs. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [6] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 26–33 vol. 1, 2005.
- [7] M. Bozorg, S. Salehkaleybar, and M. Hashemi. Seedless graph matching via tail of degree distribution for correlated Erdős-Rényi graphs. Preprint, arXiv:1907.06334.

- [8] J. A. Cain, P. Sanders, and N. Wormald. The random graph threshold for k -orientability and a fast algorithm for optimal multiple-choice allocation. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 469–476. ACM, New York, 2007.
- [9] S. Chen, S. Jiang, Z. Ma, G. P. Nolan, and B. Zhu. One-way matching of datasets with low rank signals. Preprint, arXiv:2204.13858.
- [10] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006.
- [11] D. Cullina and N. Kiyavash. Exact alignment recovery for correlated Erdos-Rényi graphs. Preprint, arXiv:1711.06783.
- [12] D. Cullina and N. Kiyavash. Improved achievability and converse bounds for erdos-renyi graph matching. In *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, SIGMETRICS '16, page 6372, New York, NY, USA, 2016. Association for Computing Machinery.
- [13] D. Cullina, N. Kiyavash, P. Mittal, and H. V. Poor. Partial recovery of Erdős-Rényi graph alignment via k -core alignment. SIGMETRICS '20, page 99100, New York, NY, USA, 2020. Association for Computing Machinery.
- [14] O. E. Dai, D. Cullina, N. Kiyavash, and M. Grossglauser. Analysis of a canonical labeling algorithm for the alignment of correlated Erdős-Rényi graphs. *Proc. ACM Meas. Anal. Comput. Syst.*, 3(2), jun 2019.
- [15] J. Ding and H. Du. Detection threshold for correlated Erdős-Rényi graphs via densest sub-graph. arXiv:2203.14573.
- [16] J. Ding, Z. Ma, Y. Wu, and J. Xu. Efficient random graph matching via degree profiles. *Probab. Theory Related Fields*, 179(1-2):29–115, 2021.
- [17] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations II: Erdős-Rényi graphs and universality. Preprint, arXiv:1907.08883.
- [18] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations: Algorithm and theory. In *Proceedings of the 37th International Conference on*

- Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2985–2995. PMLR, 13–18 Jul 2020.
- [19] S. Feizi, G. Quon, M. Medard, M. Kellis, and A. Jadbabaie. Spectral alignment of networks. Preprint, arXiv:1602.04181.
 - [20] D. Fernholz and V. Ramachandran. The k -orientability thresholds for $G_{n,p}$. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 459–468. ACM, New York, 2007.
 - [21] N. Fountoulakis, M. Khosla, and K. Panagiotou. The multiple-orientability thresholds for random hypergraphs. *Combin. Probab. Comput.*, 25(6):870–908, 2016.
 - [22] A. Frieze and M. Karoński. *Introduction to random graphs*. available at <https://www.math.cmu.edu/~af1p/BOOK.pdf>.
 - [23] L. Ganassali and L. Massoulié. From tree matching to sparse graph alignment. In J. Abernethy and S. Agarwal, editors, *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1633–1665. PMLR, 09–12 Jul 2020.
 - [24] P. Gao and N. C. Wormald. Load balancing and orientability thresholds for random hypergraphs [extended abstract]. In *STOC’10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*, pages 97–103. ACM, New York, 2010.
 - [25] B. Hajek. Performance of global load balancing by local adjustment. *IEEE Trans. Inform. Theory*, 36(6):1398–1414, 1990.
 - [26] G. Hall and L. Massoulié. Partial recovery in the graph alignment problem. Preprint, arXiv:2007.00533.
 - [27] E. Kazemi, S. H. Hassani, and M. Grossglauser. Growing a graph matching from a handful of seeds. *Proc. VLDB Endow.*, 8(10):10101021, jun 2015.
 - [28] V. Lyzinski, D. E. Fishkind, and C. E. Priebe. Seeded graph matching for correlated Erdos-Rényi graphs. *J. Mach. Learn. Res.*, 15:3513–3540, 2014.
 - [29] C. Mao, M. Rudelson, and K. Tikhomirov. Exact matching of random graphs with constant correlation. Preprint, arXiv:2110.05000.

- [30] C. Mao, Y. Wu, J. Xu, and S. H. Yu. Testing network correlation efficiently via counting trees. Preprint, arXiv:2110.11816.
- [31] E. Mossel and J. Xu. Seeded graph matching via large neighborhood statistics. *Random Structures Algorithms*, 57(3):570–611, 2020.
- [32] A. Narayanan and V. Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, 2008.
- [33] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187, 2009.
- [34] P. Pedarsani and M. Grossglauser. On the privacy of anonymized networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, page 12351243, New York, NY, USA, 2011. Association for Computing Machinery.
- [35] M. Z. Racz and A. Sridhar. Correlated randomly growing graphs. to appear in *Ann. Appl. Probab.*
- [36] M. Z. Racz and A. Sridhar. Correlated stochastic block models: Exact graph matching with applications to recovering communities. In *Advances in Neural Information Processing Systems*, 2021.
- [37] F. Shirani, S. Garg, and E. Erkip. Seeded graph matching: Efficient algorithms and theoretical guarantees. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 253–257, 2017.
- [38] R. Singh, J. Xu, and B. Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences of the United States of America*, 105:12763–8, 10 2008.
- [39] J. T. Vogelstein, J. M. Conroy, V. Lyzinski, L. J. Podrazik, S. G. Kratzer, E. T. Harley, D. E. Fishkind, R. J. Vogelstein, and C. E. Priebe. Fast approximate quadratic programming for graph matching. *PLOS ONE*, 10(4):1–17, 04 2015.
- [40] H. Wang, Y. Wu, J. Xu, and I. Yolcu. Random graph matching in geometric models: the case of complete graphs. Preprint, arXiv:2202.10662.

- [41] Y. Wu, J. Xu, and S. H. Yu. Settling the sharp reconstruction thresholds of random graph matching. Preprint, arXiv:2102.00082.
- [42] Y. Wu, J. Xu, and S. H. Yu. Testing correlation of unlabeled random graphs. Preprint, arXiv:2008.10097.
- [43] L. Yartseva and M. Grossglauser. On the performance of percolation graph matching. In *Proceedings of the First ACM Conference on Online Social Networks, COSN '13*, page 119130, New York, NY, USA, 2013. Association for Computing Machinery.

4 A Polynomial-time approximation scheme for the maximal overlap of two independent Erdős-Rényi graphs

For two independent Erdős-Rényi graphs $\mathbf{G}(n, p)$, we study the maximal overlap (i.e., the number of common edges) of these two graphs over all possible vertex correspondence. We present a polynomial-time algorithm which finds a vertex correspondence whose overlap approximates the maximal overlap up to a multiplicative factor that is arbitrarily close to 1. As a by-product, we prove that the maximal overlap is asymptotically $\frac{n}{2\alpha-1}$ for $p = n^{-\alpha}$ with some constant $\alpha \in (1/2, 1)$. This section is based on a joint work with Jian Ding and Shuyang Gong.

4.1 Introduction

In this paper we study the random optimization problem of maximizing the overlap for two independent Erdős-Rényi graphs over all possible vertex correspondence. More precisely, we fix two vertex sets $V = \{v_1, \dots, v_n\}, \mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ with edge sets

$$E_0 = \{(v_i, v_j) : 1 \leq i < j \leq n\}, \mathbf{E}_0 = \{(\mathbf{v}_i, \mathbf{v}_j) : 1 \leq i < j \leq n\}. \quad (4.1)$$

For a parameter $p \in (0, 1)$, we consider two Erdős-Rényi random graphs $G = (V, E)$ and $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ where $E \subset E_0$ and $\mathbf{E} \subset \mathbf{E}_0$ are obtained from keeping each edge in E_0 and \mathbf{E}_0 (respectively) with probability p independently. For convenience, we abuse the notation and denote by G and \mathbf{G} the adjacency matrices for the two graphs respectively. That is to say, $G_{i,j} = 1$ if and only if $(v_i, v_j) \in E$ and $\mathbf{G}_{i,j} = 1$ if and only if $(\mathbf{v}_i, \mathbf{v}_j) \in \mathbf{E}$ for all $1 \leq i < j \leq n$. Let S_n be the collection of all permutations on $[n] = \{1, \dots, n\}$. For $\pi \in S_n$, define

$$\text{Overlap}(\pi) \stackrel{\text{def}}{=} \sum_{1 \leq i < j \leq n} G_{i,j} \mathbf{G}_{\pi(i), \pi(j)} \quad (4.2)$$

to be the number of edges $(v_i, v_j) \in E$ with $(\mathbf{v}_{\pi(i)}, \mathbf{v}_{\pi(j)}) \in \mathbf{E}$. Our main result is a polynomial-time approximation scheme (PTAS) for this optimization problem.

Theorem 4.1. *Assume $p = n^{-\alpha}$ for some $\alpha \in (1/2, 1)$. For any fixed $\varepsilon > 0$, there exist a constant $C = C(\varepsilon)$ and an algorithm (see Algorithm 1) with running time $O(n^C)$ which takes G, \mathbf{G} as input and outputs a permutation $\pi^* \in S_n$, such that*

$$\mathbb{P} \left[\frac{\text{Overlap}(\pi^*)}{n} \geq \frac{1-\varepsilon}{2\alpha-1} \right] = 1 - o(1), \quad \text{as } n \rightarrow \infty. \quad (4.3)$$

In addition, $\max_{\pi \in S_n} \frac{\text{Overlap}(\pi)}{n}$ converges to $\frac{1}{2\alpha-1}$ in probability as $n \rightarrow \infty$.

Remark 4.2. For $\alpha \in (0, 1/2)$, a straightforward computation yields that with probability tending to 1 as $n \rightarrow \infty$, the overlap is asymptotic to $\frac{n^2 p^2}{2}$ for all $\pi \in S_n$. Thus, this regime is trivial if one’s goal is to approximate the maximal overlap in the asymptotic sense. For the “critical” case when α is near $1/2$, the problem seems more delicate and our current method falls short in approximating the maximal overlap asymptotically.

Remark 4.3. One can find Hamiltonian cycles in both graphs with an efficient algorithm (see [35] as well as references therein) and obviously this leads to an overlap with n edges. Also, a simple first moment computation as in the proof of Proposition 4.5 shows that the maximal overlap is of order n . Therefore, the whole challenge in this paper is to nail down the asymptotic constant $1/(2\alpha - 1)$.

Remark 4.4. Theorem 4.1 provides an *algorithmic* lower bound on the maximal overlap which is sharp asymptotically (i.e., the correct order with the correct leading constant), and we emphasize that prior to our work the asymptotic maximal overlap was not known even from a non-constructive perspective. In fact, the problem of the asymptotic maximal overlap was proposed by Yihong Wu and Jiaming Xu as an intermediate step in analyzing the hardness of matching correlated random graphs, and Yihong Wu and Jiaming Xu have discussed extensively with us on this problem. It is fair to say that we have come up with a fairly convincing roadmap using the second moment method, but the technical obstacles seem at least somewhat daunting such that we did not manage to complete the proof. The current paper was initiated with the original goal of demonstrating an *information-computation gap* for this random optimization problem (as we have believed back then), but it turned out to evolve in a somewhat unexpected way: not only the maximal overlap can be approximated by a polynomial time algorithm, but also it seems (as of now) this may even be a more tractable approach even if one’s goal is just to derive the asymptotic maximal overlap. Finally, the very stimulating discussions we had with Yihong Wu and Jiaming Xu, while not explicitly used in the current paper, have been hugely inspiring to us. We record this short “story” here as on the one hand it may be somewhat enlightening to the reader, and on the other hand we would like to take this opportunity to thank Yihong Wu and Jiaming Xu in the warmest terms.

4.1.1 Background and related results

Our motivation is twofold: on the one hand, we wish to accumulate insights for understanding the computational phase transition for matching two correlated random graphs; on the other hand, as mentioned in Remark 4.4 we wished to study the computational transition for the random optimization problem of the maximal overlap, which is a natural and important combinatorial optimization problem known to be NP-hard to solve in the worst-case scenario. We will further elaborate both of these two points in what follows.

Recently, there has been extensive study on the problem of matching the vertex correspondence between two correlated graphs and the closely related problem of detecting the correlation between two graphs. From the applied perspective, some versions of graph matching problems have emerged from various applied fields such as social network analysis [30, 31], computer vision [9, 3], computational biology [37, 38] and natural language processing [20]. From the theoretical perspective, graph matching problems seem to provide another important set of examples with the intriguing *information-computation gap*. The study of information-computation gaps for high-dimensional statistical problems as well as random optimization problems is currently an active and challenging topic. For instance, there is a huge literature on the problem of recovering communities in stochastic block models (see the monograph [1] for an excellent account on this topic) together with some related progress on the optimization problem of extremal cuts for random graphs [11, 31]; there is also a huge literature on the hidden clique problem and the closely related problem of submatrix detection (see e.g. [40] for a survey and see e.g. [18, 27] and references therein for a more or less up-to-date review). In addition, it is worth mentioning that in the last few decades various framework has been proposed to provide evidence on computational hardness for such problems with random inputs; see surveys [43, 6, 33, 24, 14] and references therein.

As for random graph matching problems, so far most of the theoretical study was carried out for Erdős-Rényi graph models since Erdős-Rényi graph is arguably the most canonical random graph model. Along this line, much progress has been made recently, including information-theoretic analysis [8, 7, 19, 45, 44, 12, 13] and proposals for various efficient algorithms [36, 46, 25, 23, 17, 38, 3, 13, 5, 9, 10, 32, 16, 20, 15, 27, 28]. As of now, it seems fair to say that we are still relatively far away from being able to completely understand the phase transition for computational complexity of graph matching problems—we believe, and we are under the impression that experts on the topic

also believe, that graph matching problems do exhibit information-computation gaps. Thus, (as a partial motivation for our study) it is plausible that understanding the computational aspects for maximizing the overlap should shed lights on the computational transition for graph matching problems.

As hinted earlier, it is somewhat unexpected that a polynomial time approximation scheme for the maximal overlap problem exists while the random graph matching problem seems to exhibit an information-computation gap. As a side remark, before finding this approximation algorithm, we have tried to demonstrate the computational hardness near the information threshold via the overlap gap property but computations suggested that this problem does not exhibit the overlap gap property. In addition, we point out that efficient algorithms have been discovered to approximate ground states for some spin glass problems [36, 26], which take advantage of the so-called full replica symmetry breaking property. From what we can tell, [36, 26] seem to be rare natural examples for which approximation algorithms were discovered for random instances whereas the worse-case problems are known to be hard to solve. In a way, our result contributes yet another example of this type and possibly of a different nature since we do not see full replica symmetry breaking in our algorithm.

4.1.2 An overview of our method

The main contribution of our paper is to propose and analyze the algorithm as in Theorem 4.1. The basic idea for our algorithm is very simple: roughly speaking, we wish to match vertices of two graphs sequentially such that the increment on the number of common edges (within matched vertices) per each matched pair is $(2\alpha - 1)^{-1}$.

An obvious issue is that $(2\alpha - 1)^{-1}$ may not be an integer. Putting this aside for a moment, we first explain our idea in the simple case when $p = n^{-\frac{3}{4} + \delta}$ with some arbitrarily small $\delta > 0$. In this case, our goal is to construct some $\pi \in S_n$ such that $\text{Overlap}(\pi) \geq (2 - \varepsilon)n$ for an arbitrarily small $\varepsilon > 0$. Assume we have already determined $\pi(1), \dots, \pi(k)$ for some $k \in [\frac{\varepsilon}{3}n, (1 - \frac{\varepsilon}{3})n]$ (i.e., assume we have already matched $v_1, \dots, v_k \in V$ to $v_{\pi(1)}, \dots, v_{\pi(k)} \in \mathbf{V}$). We wish to find some $\ell \in [n] \setminus \{\pi(1), \dots, \pi(k)\}$ such that

$$\sum_{1 \leq j \leq k} G_{j, k+1} G_{\pi(j), \ell} \geq 2. \quad (4.4)$$

Provided that this is feasible, we may set $\pi(k+1) = \ell$. Therefore, the crux is to show that (4.4)

is feasible for most of the steps. Ignoring the correlations between different steps for now, we can then regard $\sum_{1 \leq j \leq k} G_{j,k+1} \mathbf{G}_{\pi(j),\ell}$ as a Binomial variable $\mathbf{B}(k, p^2)$ for each $\ell \in [n] \setminus \{\pi(1), \dots, \pi(k)\}$ and thus (4.4) holds with probability of order $(kp^2)^2 \gtrsim n^{\delta-1}$. Since there are at least εn potential choices for ℓ , (ignoring the potential correlations for now again) it should hold with overwhelming probability that there exists at least one ℓ satisfying (4.4). This completes the heuristics underlying the success of such a simple iterative algorithm.

The main technical contribution of this work is to address the two issues we ignored so far: the integer issue and the correlations between iterations. In order to address the integer issue, it seems necessary to match χ vertices every step with an increment of ζ common edges for some suitably chosen χ, ζ with $\zeta/\chi \approx (2\alpha - 1)^{-1}$. Thus, one might think that a natural analogue of (4.4) shall be the following: there exist $\ell_1, \dots, \ell_\chi \in [n] \setminus \{\pi(1), \dots, \pi(k)\}$ such that

$$\sum_{1 \leq j \leq k} \sum_{1 \leq i \leq \chi} G_{j,k+i} \mathbf{G}_{\pi(j),\ell_i} \geq \zeta. \quad (4.5)$$

However, (4.5) is not really feasible since in order for (4.5) to hold it is necessary that for some $i \in \{1, \dots, \chi\}$, there exists $\ell \in [n] \setminus \{\pi(1), \dots, \pi(k)\}$ such that $\sum_{1 \leq j \leq k} G_{j,k+i} \mathbf{G}_{\pi(j),\ell} \geq \lceil (2\alpha - 1)^{-1} \rceil$ (where $\lceil x \rceil$ denotes the minimal integer that is at least x). A simple first moment computation suggests that the preceding requirement cannot be satisfied for most steps.

In light of the above discussions, we see that in addition to common edges between the matched vertices and the “new” vertices, it will be important to also take advantage of common edges that are within “new” vertices, which requires to also carefully choose a collection of vertices from V (otherwise typically there will be no edges within a constant number of fixed vertices). In order to implement this, it turns out that we may carefully construct a rooted tree \mathbf{T} with χ non-leaf vertices and ζ edges, and we wish that the collection of added common edges per step contains an isomorphic copy of \mathbf{T} . To be more precise, let $M = \{j \in [n] : v_j \text{ has been matched}\}$ and let $\pi(M) = \{\pi(j) : j \in M\}$, and in addition let k be the minimal integer in $[n] \setminus M$. We then wish to strengthen (4.5) to the following: there exists $\xi = \zeta + 1 - \chi$ integers $i_1, \dots, i_\xi \in M$, χ integers $j_1 = k, j_2, \dots, j_\chi \in [n] \setminus M$ and χ integers $\ell_1, \dots, \ell_\chi \in [n] \setminus \{\pi(1), \dots, \pi(k)\}$ such that the following hold:

- the subgraph of G induced on $\{v_{i_1}, \dots, v_{i_\xi}\} \cup \{v_{j_1}, \dots, v_{j_\chi}\}$ contains \mathbf{T} as a subgraph with leaf nodes $\{v_{i_1}, \dots, v_{i_\xi}\}$;

- the subgraph of G induced on $\{v_{\pi(i_1)}, \dots, v_{\pi(i_\xi)}\} \cup \{v_{\ell_1}, \dots, v_{\ell_\chi}\}$ contains \mathbf{T} as a subgraph with leaf nodes $\{v_{\pi(i_1)}, \dots, v_{\pi(i_\xi)}\}$.

(In this case, we then set $\pi(j_i) = \ell_i$ for $1 \leq i \leq \chi$.) In the above, we require the non-leaf vertices of isomorphic copies of \mathbf{T} to be contained in “new” vertices in order to make sure that the edges we add per each iteration are disjoint. In order for the existence of isomorphic copies of \mathbf{T} , we need to pose some carefully chosen balanced conditions on \mathbf{T} ; see Lemma 4.6, (4.9) and (4.10) below.

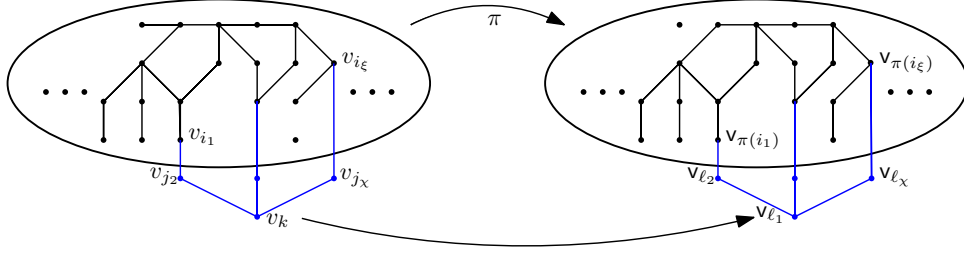


Figure 4.1: A single step in the iterative matching algorithm: the black vertices are matched ones, and the blue vertices are the ones matched in the current step (where a tree is added on both sides with leaves being black vertices).

The existence of \mathbf{T} , once appropriate balanced conditions are posed, is not that hard to show in light of [3, 28]. However, the major challenge comes in since we need to yet treat the correlations between different iterative steps. This is indeed fairly delicate and most of the difficult arguments in this paper (see Section 4.3) are devoted to address this challenge.

Finally, we point out that in order for a high precision approximation of $(2\alpha - 1)^{-1}$, we will have to choose integers χ, ζ fairly large. Since our algorithm is required to check all χ -tuples over unmatched vertices, we only obtain an algorithm with polynomial running time where the power may grow in the precision of approximation. It remains an interesting question whether a polynomial time algorithm with a fixed power can find a matching that approximates the maximal overlap up to a constant that is arbitrarily close to 1.

4.2 Proof of Theorem 4.1

4.2.1 Upper bound of the maximal overlap

In this subsection, we prove the upper bound of Theorem 4.1 in the next proposition.

Proposition 4.5. *For any constant $\rho > \frac{1}{2\alpha-1}$, we have*

$$\mathbb{P} \left[\max_{\pi \in S_n} \text{Overlap}(\pi) \geq \rho n \right] \rightarrow 0, \quad \text{as } n \rightarrow \infty. \quad (4.6)$$

Proof. By a union bound we see the left hand side of (4.6) is bounded by

$$\sum_{\pi \in S_n} \mathbb{P} [\text{Overlap}(\pi) \geq \rho n] = n! \mathbb{P} [\mathbf{B} \geq \rho n],$$

where \mathbf{B} is distributed as a binomial variable $\mathbf{B}(\binom{n}{2}, p^2)$. By standard large deviation estimates for binomial variables (see e.g. [29, Theorem 4.4]), we have

$$\mathbb{P} [\mathbf{B} \geq \rho n] \leq \exp \left(-\rho n \log \left(\frac{\rho n}{\binom{n}{2} p^2} \right) + \rho n \right) = \exp (-\rho(2\alpha - 1 + o(1))n \log n + O(n)),$$

which is of smaller magnitude than $\exp(-n \log n) \ll (n!)^{-1}$. This completes the proof. \square

We introduce some asymptotic notation which will be used later. For non-negative sequences f_n and g_n , we write $f_n \lesssim g_n$ if there exists a constant $C > 0$ such that $f_n \leq C g_n$ for all $n \geq 1$. We write $f_n \asymp g_n$ if $f_n \lesssim g_n$ and $g_n \lesssim f_n$.

4.2.2 Preliminaries of the algorithm

The rest of the paper is devoted the description and analysis of the algorithm in Theorem 4.1. Since $\alpha \in (1/2, 1)$, we may pick an integer $k \geq 1$ such that $\frac{1}{2\alpha-1} \in (k, k+1]$. Henceforth we will fix $\varepsilon > 0$ together with a positive constant η sufficiently close to 0. More precisely, we choose $\eta > 0$ such that

$$\frac{1-2\eta}{2\alpha+2\eta-1} > \frac{1-\varepsilon}{2\alpha-1}, \quad \frac{1}{2\alpha+2\eta-1} > k, \quad (4.7)$$

$$\text{and } \left\lfloor \frac{(k+1)(2(\alpha+\eta)-1)-1}{2-2(\alpha+\eta)} \right\rfloor = \left\lfloor \frac{(k+1)(2\alpha-1)-1}{2-2\alpha} \right\rfloor, \quad (4.8)$$

where $\lfloor x \rfloor$ is the greatest integer $\leq x$ and $\lceil x \rceil$ is the minimal integer $\geq x$.

As described in Section 4.1.2, our matching algorithm works by searching for isomorphic copies of certain well-chosen tree in an iterative manner. We next construct this tree with a number of “balanced conditions” which will play an essential role in the analysis of the algorithm later. Before doing so, we introduce some notation conventions. For any simple graph \mathbf{H} , let $V(\mathbf{H}), E(\mathbf{H})$ be the set of vertices and edges of \mathbf{H} , respectively. Throughout the paper, we will use bold font for a graph (e.g., \mathbf{T}) to emphasize its structural information; we use normal font for subgraphs of G (e.g., T) and we use mathsf font for subgraphs of \mathbf{G} (e.g., \mathbf{T}). In addition, we use sans-serif font such as \mathbf{L}, \mathbf{Q} to denote the subsets of $\{1, \dots, n\}$ which serve as the subscripts of vertices in G and \mathbf{G} .

Lemma 4.6. *There exist an irrational number α_η and two integers χ, ζ such that there is a rooted tree \mathbf{T} with leaf set \mathbf{L} and non-leaf set \mathbf{Q} such that the following hold:*

- (i) $\alpha < \alpha_\eta < \alpha + \eta$, $|\mathbf{Q}| = \chi$, $|E(\mathbf{T})| = \zeta$ and $0 < \chi - (2\alpha_\eta - 1)\zeta < 1 - \alpha_\eta$.
- (ii) For each $u \in \mathbf{Q}$ which is adjacent to some leaves, u is adjacent to exactly k leaves.
- (iii) For any subgraph $\mathbf{F} \subsetneq \mathbf{T}$ with $\mathbf{L} \subset \mathbf{F}$, it holds that

$$|V(\mathbf{T}) \setminus V(\mathbf{F})| < \alpha_\eta |E(\mathbf{T}) \setminus E(\mathbf{F})|. \quad (4.9)$$

- (iv) For any subtree $\mathbf{T}_0 \subsetneq \mathbf{T}$, it holds that

$$|V(\mathbf{T}_0) \cap \mathbf{Q}| - (2\alpha_\eta - 1)|E(\mathbf{T}_0)| > \chi - (2\alpha_\eta - 1)\zeta. \quad (4.10)$$

Proof. Recall that $k = \lceil \frac{1}{2\alpha - 1} \rceil - 1$. First, we claim there exist positive integers $n_1 \leq n_2 \leq \dots \leq n_l$ such that

$$2\alpha - 1 < \frac{1 + \frac{1}{n_1} + \frac{1}{n_1 n_2} + \dots + \frac{1}{n_1 n_2 \dots n_l}}{k + 1 + \frac{1}{n_1} + \frac{1}{n_1 n_2} + \dots + \frac{1}{n_1 n_2 \dots n_{l-1}}} < 2(\alpha + \eta) - 1. \quad (4.11)$$

In order to see this, pick an *irrational* number β such that

$$2\alpha - 1 < \frac{1 + \beta}{k + 1 + \beta} < 2(\alpha + \eta) - 1.$$

Then we can express β as the infinite sum

$$\beta = \frac{1}{n_1} + \frac{1}{n_1 n_2} + \frac{1}{n_1 n_2 n_3} + \dots$$

with positive integers $n_1 \leq n_2 \leq \dots$ determined by the following inductive procedure: assuming n_1, \dots, n_{k-1} have been fixed, we choose n_k such that

$$\frac{1}{n_k} < n_1 \dots n_{k-1} \left(\beta - \frac{1}{n_1} - \dots - \frac{1}{n_1 \dots n_{k-1}} \right) < \frac{1}{n_k - 1}$$

(note that the irrationality of β ensures the strict inequality above). It is straightforward to verify that this yields the desired expression. With this at hand, we obtain the desired relation (4.11) by an appropriate truncation.

Writing $\ell = n_1 n_2 \cdots n_l$, we set

$$\chi = \ell \left(1 + \frac{1}{n_1} + \cdots + \frac{1}{n_1 n_2 \cdots n_l} \right) \text{ and } \zeta = \ell \left(k + 1 + \frac{1}{n_1} + \cdots + \frac{1}{n_1 n_2 \cdots n_{l-1}} \right).$$

Thus, we have $\zeta = k\ell + \chi - 1$. We define a rooted tree \mathbf{T} with $(\ell + 2)$ generations as follows: the 0-th generation contains a single root; for $0 \leq i \leq l - 1$, each vertex in the i -th generation has n_{l-i} children; each vertex in the l -th generation has k children (which are leaves). It is clear that \mathbf{T} has χ non-leaf vertices and ζ edges. In addition, \mathbf{T} satisfies (ii). Choose $\tilde{\alpha}$ such that $\chi/\zeta = 2\tilde{\alpha} - 1$. Then by (4.11), we see that (i) holds for $\alpha_\eta \in (\alpha, \tilde{\alpha})$ which are sufficiently close to $\tilde{\alpha}$. In the rest of proof, we will show that for some irrational $\alpha_\eta < \tilde{\alpha}$ sufficiently close to $\tilde{\alpha}$, (iii) and (iv) also hold. We remark that the proof is somewhat technical and it may be skipped at the first reading.

We begin with (4.9). We say \mathbf{T}_0 is an *entire subtree* of \mathbf{T} if \mathbf{T}_0 is a subtree of \mathbf{T} such that \mathbf{T}_0 contains all neighbors of u in \mathbf{T} for any vertex u with degree larger than 1 in \mathbf{T}_0 . For a subgraph $\mathbf{F} \subset \mathbf{T}$, we consider the following procedure:

- remove all the vertices and edges in \mathbf{F} from \mathbf{T} and thus what remains is a union of vertices and edges while some edges may be incomplete since its one or both endpoints may have been removed;
- for each aforementioned incomplete edge that remains and for each of the endpoint that has been removed, we add a *distinct* vertex to replace the removed endpoint.

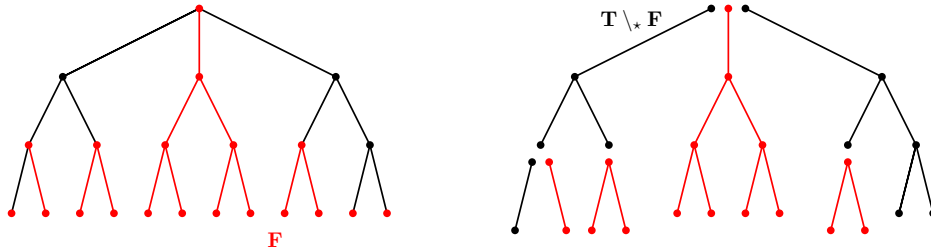


Figure 4.2: The red part on the left is the subgraph \mathbf{F} . The black part on the right is $\mathbf{T} \setminus_\star \mathbf{F}$. There are three entire subtrees (black) in $\mathbf{T} \setminus_\star \mathbf{F}$.

At the end of this procedure, we obtain a collection of trees that are both edge-disjoint and vertex-disjoint and we denote this collection of trees as $\mathbf{T} \setminus_\star \mathbf{F}$. In addition, for each tree in $\mathbf{T} \setminus_\star \mathbf{F}$, it can be regarded as a subtree of \mathbf{T} (if we identify the added vertex in the tree in $\mathbf{T} \setminus_\star \mathbf{F}$ to the corresponding vertex in \mathbf{T}) and one can verify that each such subtree is an entire subtree of \mathbf{T} .

Therefore, in order to prove (4.9) it suffices to show that for any entire subtree $\mathbf{T}_0 \subset \mathbf{T}$ where \mathbf{Q}_0 is the collection of vertices with degree larger than 1 in \mathbf{T}_0 , we have

$$|\mathbf{Q}_0| \leq \alpha |E(\mathbf{T}_0)| \quad (4.12)$$

(note that $\alpha < \alpha_\eta$) and in addition the equality holds only when the tree is a singleton. For a subtree \mathbf{T}_0 of \mathbf{T} , assume its root \mathbf{v} (i.e., the vertex in $V(\mathbf{T}_0)$ closest to the root of \mathbf{T}) lies in the i -th generation of \mathbf{T} . We define the height of \mathbf{T}_0 by

$$h(\mathbf{T}_0) = \begin{cases} l + 1 - i, & \text{if } |V(\mathbf{T}_0)| \leq 2 \text{ or } \mathbf{v} \in \mathbf{Q}_0, \\ l - i, & \text{if } |V(\mathbf{T}_0)| > 2 \text{ and } \mathbf{v} \notin \mathbf{Q}_0. \end{cases}$$

Since obviously the equality in (4.12) holds if $|V(\mathbf{T}_0)| \leq 2$, we then deal with the other cases.

Let $m = |\{i : n_i = 1\}|$. From (4.11) we see

$$\frac{1}{n_1} + \cdots + \frac{1}{n_1 \cdots n_{l-1}} < \frac{(k+1)(2\tilde{\alpha}-1)-1}{2-2\tilde{\alpha}} < \frac{1}{n_1} + \cdots + \frac{1}{n_1 \cdots n_l}.$$

Since $n_1 = \cdots = n_m = 1$ and $n_t \geq 2$ for $t > m$, we get from (4.8) that

$$m = \left\lfloor \frac{(k+1)(2\tilde{\alpha}-1)-1}{2-2\tilde{\alpha}} \right\rfloor = \left\lfloor \frac{(k+1)(2\alpha-1)-1}{2-2\alpha} \right\rfloor. \quad (4.13)$$

We now prove (4.12) for any entire subtree \mathbf{T}_0 with $h(\mathbf{T}_0) \leq m+1$.

- If $1/2 < \alpha \leq 3/4$, then we have $k \geq 2$ and $m = 0$. Thus, \mathbf{T}_0 must be a star graph with k or $k+1$ leaves, implying that

$$|\mathbf{Q}_0| - \alpha |E(\mathbf{T}_0)| \leq 1 - k\alpha < 0.$$

- If $3/4 < \alpha < 1$, then $k = 1$ and $\frac{2m+3}{2m+4} < \alpha \leq \frac{2m+5}{2m+6}$. Thus, \mathbf{T}_0 must be a chain with length at most $m+2$, implying that

$$|\mathbf{Q}_0| - \alpha |E(\mathbf{T}_0)| \leq (m+1) - (m+2)\alpha < 0.$$

For an entire subtree \mathbf{T}_0 with $h(\mathbf{T}_0) \geq m + 2$, we prove (4.12) by proving it for generalized entire subtrees, where \mathbf{T}_0 is a generalized entire subtree if \mathbf{T}_0 contains all the children (in \mathbf{T}) of u for each $u \in \mathbf{Q}_0$. We prove a stronger version of (4.12) since we will prove it by induction (and it is not uncommon that proving a stronger statement by induction makes the proof easier). For the base case when $h(\mathbf{T}_0) = m + 2$, it follows since

$$|\mathbf{Q}_0| - \alpha|E(\mathbf{T}_0)| \leq \begin{cases} 1 - 2\alpha, & 1/2 < \alpha \leq 3/4, \\ (2m + 3) - (2m + 4)\alpha, & 3/4 < \alpha < 1, \end{cases}$$

which is negative by preceding discussions. Now suppose that (4.12) holds whenever \mathbf{T}_0 is a generalized entire subtree with $h(\mathbf{T}_0) = h$, and we next consider a generalized entire subtree \mathbf{T}_0 with $h(\mathbf{T}_0) = h + 1$. Denote the non-leaf vertex of \mathbf{T}_0 at the $(l - h - 1)$ -th generation (in \mathbf{T}) by \mathbf{o}_0 (this is well-defined since \mathbf{o}_0 uniquely exists as long as \mathbf{T}_0 contains more than 2 vertices). In addition, denote the subtrees of \mathbf{T}_0 rooted at children of \mathbf{o}_0 by $\mathbf{T}_1, \dots, \mathbf{T}_s$ where \mathbf{T}_i contains all the descendants (in \mathbf{T}_0) of its root. Then \mathbf{T}_i is a generalized entire subtree with height h for $i = 1, \dots, s$. Thus, by our induction hypothesis we get (denoting by \mathbf{Q}_i the non-leaf vertices in \mathbf{T}_i)

$$|\mathbf{Q}_i| \leq \alpha|E(\mathbf{T}_i)| \text{ for all } 1 \leq i \leq s. \quad (4.14)$$

Note that $|\mathbf{Q}_0| = 1 + |\mathbf{Q}_1| + \dots + |\mathbf{Q}_s|$ and $|E(\mathbf{T}_0)| \geq s + |E(\mathbf{T}_1)| + \dots + |E(\mathbf{T}_s)|$. Combined with (4.14) and the fact $s \geq 2$, it yields that $|\mathbf{Q}_0| - \alpha|E(\mathbf{T}_0)| < 0$, completing the inductive step for the proof of (4.12) (so this proves (4.9)).

We next prove (4.10). It suffices to show that for any subtree $\mathbf{T}_0 \subsetneq \mathbf{T}$,

$$\frac{|V(\mathbf{T}_0) \cap \mathbf{Q}|}{|E(\mathbf{T}_0)|} > \frac{|\mathbf{Q}|}{|E(\mathbf{T})|} = \frac{\chi}{\zeta} = 2\tilde{\alpha} - 1. \quad (4.15)$$

Indeed, (4.15) implies that

$$|V(\mathbf{T}_0) \cap \mathbf{Q}| - (2\tilde{\alpha} - 1)|E(\mathbf{T}_0)| > 0 = \chi - (2\tilde{\alpha} - 1)\zeta \text{ for all } \mathbf{T}_0 \subsetneq \mathbf{T}.$$

Writing A the set of α_η for which (4.10) holds, we get from the preceding inequality that $\tilde{\alpha} \in A$. By the finiteness of \mathbf{T} , we have that A is an open set and thus we can pick an irrational $\alpha_\eta \in A$ with $\alpha_\eta < \tilde{\alpha}$ and sufficiently close to $\tilde{\alpha}$.

Define $\mathbb{T} = \{\mathbf{T}_v : v \in \mathbf{T}\}$ where \mathbf{T}_v is a subtree rooted at v containing all descendants of v in \mathbf{T} . We first verify (4.15) for subtrees in \mathbb{T} . Denoting $n_0 = 1$, we see that for any $0 \leq j \leq l$ and any

vertex v in the $(l-j)$ -th generation,

$$\frac{|V(\mathbf{T}_v) \cap \mathbf{Q}|}{|E(\mathbf{T}_v)|} = \frac{\frac{1}{n_0} + \frac{1}{n_1} + \cdots + \frac{1}{n_1 \cdots n_j}}{k + \frac{1}{n_0} + \frac{1}{n_1} + \cdots + \frac{1}{n_1 \cdots n_{j-1}}} \stackrel{\text{def}}{=} 2\alpha_j - 1.$$

We now show $\alpha_j > \tilde{\alpha}$ for any $0 \leq j \leq l-1$. This is true for $j=0$ by the choice of η in (4.7). For $j \geq 1$, we assume otherwise that the inequality does not hold for some $j \geq 1$. Then we may pick the minimal $j_0 \in \{1, 2, \dots, l-1\}$ such that $\alpha_{j_0} \leq \tilde{\alpha}$. By minimality of j_0 , we have that $\alpha_{j_0-1} > \tilde{\alpha}$ and thus $n_{j_0} > \frac{1}{2\tilde{\alpha}-1}$. Since $\{n_j\}_{j=0}^l$ is non-decreasing, we get $n_j > \frac{1}{2\tilde{\alpha}-1}$ for all $j \geq j_0$, and thus $\alpha_j < \tilde{\alpha}$ for any $j > j_0$. This implies $\tilde{\alpha} = \alpha_l < \tilde{\alpha}$, arriving at a contradiction and thus verifying (4.15) for subtrees in \mathbb{T} .

Finally, we reduce the general case to subtrees in \mathbb{T} as follows: for any subtree $\mathbf{T}_0 \subset \mathbf{T}$, we denote its root by \mathbf{o}_0 (this is the closest vertex in \mathbf{T}_0 to the root of \mathbf{T}) and consider $\mathbf{T}_{\mathbf{o}_0} \in \mathbb{T}$. It is clear that \mathbf{T}_0 can be obtained from $\mathbf{T}_{\mathbf{o}_0}$ by deleting some vertices and edges. Further, the numbers of vertices and edges deleted are the same and let us denote by N this number. We then have

$$\frac{|V(\mathbf{T}_0) \cap \mathbf{Q}|}{|E(\mathbf{T}_0)|} \geq \frac{|V(\mathbf{T}_{\mathbf{o}_0}) \cap \mathbf{Q}| - N}{|E(\mathbf{T}_{\mathbf{o}_0})| - N} \geq \frac{|\mathbf{T}_{\mathbf{o}_0} \cap \mathbf{Q}|}{E(\mathbf{T}_{\mathbf{o}_0})} \geq 2\tilde{\alpha} - 1, \quad (4.16)$$

with equality holds if and only if $N=0$ and $\mathbf{T}_{\mathbf{o}_0} = \mathbf{T}$ (that is, $\mathbf{T}_0 = \mathbf{T}$). This completes the proof of (4.15). \square

In what follows, we fix α_η, χ, ζ and the tree \mathbf{T} given in Lemma 4.6, and let $\xi = \zeta - \chi + 1$. We label the vertices of \mathbf{T} by $1, 2, \dots, \chi, \chi+1, \dots, \chi+\xi$, such that the root is labeled by 1, $\mathbf{Q} = \{1, 2, \dots, \chi\}$ and $\mathbf{L} = \{\chi+1, \dots, \chi+\xi\}$. Recall that

$$E(\mathbf{T}) = \{(i, j) : 1 \leq i < j \leq \chi + \xi, i \text{ is adjacent to } j \text{ in } \mathbf{T}\}. \quad (4.17)$$

For later proof, it would be convenient to consider Erdős-Rényi graphs with edge density $p_\eta = n^{-\alpha_\eta}$, and we can reduce our problem to this case by monotonicity. Indeed, since $\alpha_\eta > \alpha$ we have $p_\eta < p$ for all n and thus $\mathbf{G}(n, p)$ is stochastically dominated by $\mathbf{G}(n, p_\eta)$. In addition, since monotonicity in p naturally holds for $\text{Overlap}(\pi)$, we may change the underlying distribution from $\mathbf{G}(n, p)$ to $\mathbf{G}(n, p_\eta)$ and prove the lower bound for the latter. In what follows we use G and \mathbf{G} to denote two independent Erdős-Rényi graphs $\mathbf{G}(n, p_\eta)$ (as well as the respective adjacency matrices); we drop the superscript η here for notation convenience. We denote by \mathbb{P} the joint law of (G, \mathbf{G}) . Finally, we remark that the assumption of irrationality for α_η in Lemma 4.6 is purely technical, which will be useful in later proofs (for the purpose of ruling out equalities).

4.2.3 Description of the algorithm

We first introduce some notations. For any nonempty subset $S \subset [n]$ and any integer $1 \leq m \leq |S|$, define (recalling (4.17))

$$\mathfrak{A}(S, m) = \{(i_1, \dots, i_m) \in S^m : i_1, \dots, i_m \text{ are distinct}\}. \quad (4.18)$$

In addition, for $m \in \{\chi, \xi\}$, we sample a random total ordering \prec_m on the set $\mathfrak{A}([n], m)$ uniformly from all possible orderings and fix it (This can be done efficiently, since it is equivalent to sampling a uniform permutation on a set with cardinality polynomial in n , which can be done in poly-time in n). We will omit the subscript when it is clear in the context. For any tuple $L = (t_1, \dots, t_m)$ and any permutation π , we write $\pi(L) = (\pi(t_1), \dots, \pi(t_m))$. For a ξ -tuple $L = (t_{\chi+1}, \dots, t_{\chi+\xi}) \in \mathfrak{A}([n], \xi)$ and a χ -tuple $Q = (t_1, \dots, t_\chi) \in \mathfrak{A}([n], \chi)$ with $L \cap Q = \emptyset$ (i.e., the coordinates of L are disjoint from the coordinates of Q), we let $L = \{v_i : i \in L\}$, $Q = \{v_i : i \in Q\}$ and define

$$\{L \bowtie_G Q\} = \{G_{t_i, t_j} = 1 \text{ for all } (i, j) \in E(\mathbf{T})\}. \quad (4.19)$$

Let $\{L \not\bowtie_G Q\}$ be the complement of $\{L \bowtie_G Q\}$. In addition, similar notations of L , Q , $\{L \bowtie_G Q\}$, $\{L \not\bowtie_G Q\}$ apply for the graph G , where the $G_{i,j}$'s in (4.19) are replaced by corresponding $G_{i,j}$'s. For any ξ -tuple $L \subset \mathfrak{A}([n], \xi)$ and any subset $U \subset [n]$, let $\{L \bowtie_G U\}$ be the event that $\{L \bowtie_G Q\}$ occurs for at least one χ -tuple $Q \subset \mathfrak{A}(U \setminus L, \chi)$ where $U = \{v_i : i \in U\}$. Further, for $r \in U$, let $\{L \bowtie_{G,r} U\}$ denote the event that there exists $Q = (t_1, \dots, t_\chi) \in \mathfrak{A}(U \setminus L, \chi)$ such that the event $\{L \bowtie_G Q\}$ holds and $t_1 = r$. Similar notations of $\{L \bowtie_G U\}$, $\{L \bowtie_{G,r} U\}$ apply for G . Moreover, we define a mapping $\Pi = \Pi_\pi : [n] \rightarrow V$ corresponding to π such that $\Pi(i) = v_{\pi(i)}$, and we define a mapping $I_G : [n] \rightarrow V$ by $I_G(i) = v_i$ (we also define I_G similarly).

Now we are ready to describe our matching algorithm. Roughly speaking, our algorithm proceeds iteratively in the following greedy sense: in the s -th step, assuming we have already matched a set M_{s-1} (i.e., we have determined the value of $\pi^*(i)$ for $i \in M_{s-1}$), we then pick some $u \in R_{s-1} = [n] \setminus M_{s-1}$ and try to find a triple of tuples (L, Q, Q) with

$$\begin{aligned} L &= \{(v_{t_{\chi+1}}, \dots, v_{t_{\chi+\xi}}) : (t_{\chi+1}, \dots, t_{\chi+\xi}) \in \mathfrak{A}(M_{s-1}, \xi)\}, \\ Q &= \{(v_{t_1}, \dots, v_{t_\chi}) : (t_1, \dots, t_\chi) \in \mathfrak{A}(R_{s-1}, \chi)\}, \\ Q &= \{(v_{t'_1}, \dots, v_{t'_\chi}) : (t'_1, \dots, t'_\chi) \in \mathfrak{A}([n] \setminus \pi^*(M_{s-1}), \chi)\}. \end{aligned}$$

such that $t_1 = u$ and both $\{L \bowtie_G Q\}$ and $\{\Pi(I_G^{-1}(L)) \bowtie_G Q\}$ happen. If such L, Q, Q exist, we let $\pi^*(t_j) = t'_j$ and $\Pi(t_i) = v_{t'_j}$ for $1 \leq j \leq \chi$ (that is, we determine the values of π^* at t_1, \dots, t_χ

in this step); else we just choose the value $\pi^*(u)$ arbitrarily (that is, we just determine the value of π^* at u in an arbitrary manner). This completes the construction for the s -th step. We expect that in most steps we are able to find such triples and thus the increment for the overlap is at least ζ . Since there are around n/χ steps in total, this would imply that at the end of the procedure $\text{Overlap}(\pi^*) \approx \zeta n/\chi \approx \frac{n}{2\alpha-1}$, as desired.

While the above algorithm may indeed achieve our goal, we will further incorporate the following technical operations in order to facilitate the analysis of the algorithm:

- Fixing a large integer κ_0 with $\kappa_0 > 4\zeta/\eta$, in each step we will first exclude the vertices which have been “used” for at least κ_0 times.
- In each step, when seeking the desired triple (L, Q, Q) , we will check the tuples in the order given by \prec .

We next present our full algorithm formally. Initially, we set $\pi^*(i) = i$ for all $1 \leq i \leq \eta n$ and let $M_0 = \{1, 2, \dots, \lfloor \eta n \rfloor\}$. Let $R_0 = [n] \setminus M_0$. Set $\text{EXP}_0 = \text{SUC}_0 = \text{FAIL}_0 = \emptyset$. As we will see in the formal definition below, we will define $\text{EXP}_s = \text{SUC}_s \cup \text{FAIL}_s \subset \mathfrak{A}([n], \xi)$ to be the set of ξ -tuples which are explored during the s -th step, where SUC_s (respectively FAIL_s) denotes the collection of tuples which were successfully matched (respectively, failed to be matched). Our algorithm then proceeds as follows:

Algorithm 1 Greedy Matching Algorithm

- 1: Define $\pi^*(i), 1 \leq i \leq \eta n$ and $M_0, R_0, \text{EXP}_0, \text{SUC}_0, \text{FAIL}_0$ as above.
- 2: **for** $s = 1, 2, \dots$ **do**
- 3: **if** $|R_{s-1}| \geq \eta n$ **then**
- 4: Set $I_s = 0$; set a triple $\text{MT}_s = \text{null}$; set $M_s = M_{s-1}$.
- 5: **for all** $u \in M_{s-1}$ **do**
- 6: **if** u appears in at least κ_0 tuples in $\{\text{MT}_t : 1 \leq t \leq s-1\}$ **then**
- 7: Delete u from M_s .
- 8: **end if**
- 9: **end for**
- 10: Set $\text{EXP}_s = \text{SUC}_s = \text{FAIL}_s = \emptyset$.
- 11: Find the minimal element $u_s \in R_{s-1}$.
- 12: Let CAND_s be the collection of $L \in \mathfrak{A}(M_s, \xi)$ so that $\{I_G(L) \bowtie_{G, u_s} R_{s-1}\}$ holds.

```

13:   Label elements in  $\text{CAND}_s$  by  $L_1, \dots, L_l$  such that  $L_1 \prec \dots \prec L_l$ .
14:   Label all  $\chi$ -tuples in  $\mathfrak{A}([n] \setminus \pi^*(M_{s-1}), \chi)$  by  $Q_1, \dots, Q_m$  so that  $Q_1 \prec \dots \prec Q_m$ .
15:   for  $i = 1, 2, \dots, l$  do
16:     for  $j = 1, 2, \dots, m$  do
17:       Check whether  $\{\Pi(L_i) \bowtie_G I_G(Q_j)\}$  happens where  $\Pi = \Pi_{\pi^*}$ .
18:       if  $\{\Pi(L_i) \bowtie_G I_G(Q_j)\}$  happens then
19:         Find a  $\chi$ -tuple  $Q \in \mathfrak{A}(R_{s-1}, \chi)$  such that  $\{I_G(L_i) \bowtie_{G, u_s} I_G(Q)\}$  holds. {The
           existence of  $Q$  is guaranteed by definition of  $L_i$ .}
20:         Set  $I_s = 1$  and  $\text{MT}_s = (L_i, Q, Q_j)$ .
21:         Add  $L_i$  into  $\text{SUC}_s$ . Then break the for cycle, and break the for cycle again.
22:       else if  $j = m$  then
23:         Add  $L_i$  into  $\text{FAIL}_s$ . {This means we have checked all possible tuples  $Q_j, j =$ 
            $1, 2, \dots, m$  but failed to find the desired one.}
24:       end if
25:     end for
26:   end for
27:   Set  $\text{EXP}_s = \text{SUC}_s \cup \text{FAIL}_s$ .
28:   if  $I_s = 1$  then
29:     Recall  $\text{MT}_s = (L_i, Q, Q_j)$ . Set  $\pi^*$  on coordinates of  $Q$  so that  $\pi^*(Q) = Q_j$ .
30:     Set  $M_s$  as the union of  $M_{s-1}$  and all coordinates in  $Q$ ; set  $R_s = [n] \setminus M_s$ .
31:   else
32:     Set  $\pi^*(u_s)$  as the minimal element in  $[n] \setminus \pi^*(M_{s-1})$ .
33:     Set  $M_s = M_{s-1} \cup \{u_s\}$ ,  $R_s = [n] \setminus M_s$ .
34:   end if
35: else
36:   set  $\pi^*(u)$  for  $u \in R_s$  arbitrarily such that  $\pi^*$  becomes a permutation on  $[n]$ .
37:   Break the for cycle.
38: end if
39: end for
40: return  $\pi^*$ .

```

Remark 4.7. CAND_s stands for the candidate tuples which may be successfully matched in the

s -th step, while EXP_s denotes for the set of tuples in CAND_s that have been checked during the algorithm. In most part of this paper, we do not distinguish CAND_s and EXP_s , and they are indeed equal when $I_s = 0$ (i.e., we have checked all candidate tuples but failed to match up). However, one should keep in mind that typically it holds $|\text{EXP}_s| \ll |\text{CAND}_s|$. Heuristically, this is because the checking procedure is according to a uniform ordering \prec and it stops as long as any successful matching triple is found. We shall make this precise and incorporate it as an ingredient for proofs in Section 4.3.3.

4.2.4 Analysis of the algorithm

In this subsection we prove that Algorithm 1 satisfies the condition of Theorem 4.1. To this end, we need to analyze the conditional probability for the event $\{I_s = 1\}$ given the behavior of Algorithm 1 in previous steps. For notation convenience, in what follows we write $\pi = \pi^*$ and write $\mathbf{R}'_s = [n] \setminus \pi(\mathbf{M}_s)$. Since in each step we always have $|\mathbf{R}_{s+1}| \geq |\mathbf{R}_s| - \chi$, the algorithm runs for at least $S = \lfloor \frac{1-2\eta}{\chi} n \rfloor$ steps. We will only consider the first S steps. For each $1 \leq s \leq S$, define \mathcal{F}_{s-1} to be the σ -field generated by $\mathbf{M}_t, \text{SUC}_t, \text{FAIL}_t, \text{CAND}_t, I_t, \text{MT}_t$ for $t = 1, \dots, s-1$ as well as $\pi(i)$ for $i \in \mathbf{M}_{s-1}$ (here we denote the matching triples in MT_t by $(\mathbf{L}_t, \mathbf{Q}_t, \mathbf{Q}'_t)$ if $I_t = 1$, and denote $\text{MT}_t = \text{null}$ if $I_t = 0$). Then \mathcal{F}_{s-1} contains all the information generated by Algorithm 1 in the first $s-1$ steps. Thus conditioning on a realization of \mathcal{F}_{s-1} is equivalent to conditioning on a realization of the first $s-1$ steps of Algorithm 1. We further denote $\mathcal{F}_{s-1/2}$ as the σ -field generated by \mathcal{F}_{s-1} and CAND_s .

With slight abuse of notation, we will write $\mathbb{P}[\cdot \mid \mathcal{F}_{s-1}]$ (respectively $\mathbb{P}[\cdot \mid \mathcal{F}_{s-1/2}]$) for the conditional probability given some particular realization of \mathcal{F}_{s-1} (respectively some realization of $\mathcal{F}_{s-1/2}$). Let

$$\begin{aligned} \text{Fail}_{s-1} = & \bigcup_{1 \leq t \leq s-1} \left\{ (\Pi(\mathbf{L}), I_G(\mathbf{Q}')) : \mathbf{L} \in \text{FAIL}_t, \mathbf{Q}' \in \mathfrak{A}(\mathbf{R}'_t, \chi) \right\} \\ & \bigcup_{\substack{1 \leq t \leq s-1, \\ I_t=1}} \left\{ (\Pi(\mathbf{L}_t), I_G(\mathbf{Q}')) : \mathbf{Q}' \in \mathfrak{A}(\mathbf{R}'_t, \chi), \mathbf{Q}' \prec \mathbf{Q}'_t \right\} \end{aligned} \quad (4.20)$$

and

$$\text{Suc}_{s-1} = \bigcup_{\substack{1 \leq t \leq s-1, \\ I_t=1}} \{(I_G(\mathbf{L}_t), I_G(\mathbf{Q}_t))\}, \quad \text{Suc}_{s-1} = \bigcup_{\substack{1 \leq t \leq s-1, \\ I_t=1}} \{(\Pi(\mathbf{L}_t), I_G(\mathbf{Q}'_t))\}. \quad (4.21)$$

Since the two graphs are independent, we have that for any event \mathcal{B} measurable with respect to \mathbf{G} ,

$$\mathbb{P}[\mathcal{B} \mid \mathcal{F}_{s-1/2}] = \mathbb{P}[\mathcal{B} \mid \mathcal{F}_{s-1}] = \mathbb{P}[\mathcal{B} \mid \mathcal{A}_s^1, \mathcal{A}_s^2]$$

with

$$\mathcal{A}_s^1 = \bigcap_{(\mathbf{L}, \mathbf{Q}) \in \text{Fail}_{s-1}} \{\mathbf{L} \bowtie_{\mathbf{G}} \mathbf{Q}\}, \quad \mathcal{A}_s^2 = \bigcap_{(\mathbf{L}, \mathbf{Q}) \in \text{Suc}_{s-1}} \{\mathbf{L} \bowtie_{\mathbf{G}} \mathbf{Q}\}. \quad (4.22)$$

Now we investigate the event $\{I_s = 0\}$ conditioned on some realization of \mathcal{F}_{s-1} . As in the algorithm, we first pick $u_s \in R_{s-1}$ and find all ξ -tuples $\mathbf{L} \in \text{CAND}_s$ (recall that $I_G(\mathbf{L}) \bowtie_{G, u_s} R_{s-1}$ for $\mathbf{L} \in \text{CAND}_s$). Note that whenever \mathcal{F}_{s-1} is given, this procedure is measurable with respect to G . Provided with CAND_s , we have that $I_s = 0$ is equivalent to

$$\sum_{\mathbf{L} \in \text{CAND}_s} \mathbf{1}_{\{\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1}\}} = 0 \text{ where } R_{s-1} = I_G(R'_{s-1}).$$

Since our aim is to lower-bound the preceding sum, it turns out more convenient to consider the sum over tuples with additional properties. To this end, we will define certain s -good tuples (s -good is measurable with respect to \mathcal{F}_{s-1} ; see Definition 4.13 below) and consider

$$X_s \stackrel{\text{def}}{=} \sum_{\substack{\mathbf{L} \in \text{CAND}_s \\ \mathbf{L} \text{ is } s\text{-good}}} \mathbf{1}_{\{\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1}\}}. \quad (4.23)$$

Clearly, $\{I_s = 0\} \subset \{X_s = 0\}$.

For any $1 \leq s \leq S$, we shall define two good events $\mathcal{G}_s^i, i = 1, 2$, where \mathcal{G}_s^1 is measurable with respect to \mathcal{F}_{s-1} and \mathcal{G}_s^2 is measurable with respect to $\mathcal{F}_{s-1/2}$. The precise definitions of $\mathcal{G}_s^1, \mathcal{G}_s^2$ will be given in the next section. Roughly speaking, \mathcal{G}_s^1 states that the cardinality of Fail_{s-1} is not unusually large, while \mathcal{G}_s^2 says that the number of s -good tuples in CAND_s is not too small and the intersecting profile for pairs of s -good tuples in CAND_s behaves in a typical way.

The following three propositions, whose proofs are postponed until Section 4.3, are key ingredients for the proof of Theorem 4.1.

Proposition 4.8. *For $1 \leq s \leq S$, any realization of \mathcal{F}_{s-1} and any ξ -tuple $\mathbf{L} \in \mathfrak{A}(\mathbf{M}_{s-1}, \xi)$,*

$$\mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1} \mid \mathcal{A}_s^1, \mathcal{A}_s^2] \leq \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1}]. \quad (4.24)$$

Furthermore, for any realization on the good event \mathcal{G}_s^1 and s -good ξ -tuple $\mathbf{L} \in \mathfrak{A}(\mathbf{M}_{s-1}, \xi)$, it holds uniformly that

$$\mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1} \mid \mathcal{A}_s^1, \mathcal{A}_s^2] \geq [1 - o(1)] \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} R_{s-1}]. \quad (4.25)$$

In addition, under such additional assumptions, for some positive constants c_1, c_2 depending only on η, α_η and \mathbf{T} , we have

$$c_1 n^\chi p_\eta^\zeta \leq \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{A}_s^1, \mathcal{A}_s^2] \leq c_2 n^\chi p_\eta^\zeta. \quad (4.26)$$

Proposition 4.9. For any realization of $\mathcal{F}_{s-1/2}$ on $\mathcal{G}_s^1 \cap \mathcal{G}_s^2$, it holds uniformly for $1 \leq s \leq S$ that

$$\mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1/2}] \leq [1 + o(1)] (\mathbb{E}[X_s \mid \mathcal{F}_{s-1/2}])^2. \quad (4.27)$$

Proposition 4.10. The good events are typical, i.e. for $1 \leq s \leq S$, it holds uniformly that

$$\mathbb{P}[\mathcal{G}_s^1] = 1 - o(1), \quad \mathbb{P}[\mathcal{G}_s^2] = 1 - o(1). \quad (4.28)$$

Proof of Theorem 4.1. It is clear that this algorithm runs in polynomial time, so it remains to show (4.3). For each $1 \leq s \leq S$, whenever the realization of $\mathcal{F}_{s-1/2}$ satisfies $\mathcal{G}_s^1 \cap \mathcal{G}_s^2$, we can apply the second moment method and obtain that

$$\mathbb{P}[I_s = 1 \mid \mathcal{F}_{s-1/2}] \geq \mathbb{P}[X_s > 0 \mid \mathcal{F}_{s-1/2}] \geq \frac{(\mathbb{E}[X_s \mid \mathcal{F}_{s-1/2}])^2}{\mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1/2}]}.$$

Thus, from Proposition 4.9 we get $\mathbb{P}[I_s = 0 \mid \mathcal{F}_{s-1/2}] = o(1)$, uniformly for all $1 \leq s \leq S$ and all realization of $\mathcal{F}_{s-1/2}$ satisfying $\mathcal{G}_s^1 \cap \mathcal{G}_s^2$. Combined with Proposition 4.10, it yields that

$$\begin{aligned} \mathbb{E}\left[\sum_{s=1}^S \mathbf{1}_{\{I_s=0\}}\right] &= \sum_{s=1}^S \left(\mathbb{E}[\mathbf{1}_{\{I_s=0\}} \cap (\mathcal{G}_s^1 \cap \mathcal{G}_s^2)^c] + \mathbb{E}[\mathbf{1}_{\{I_s=0\}} \cap \mathcal{G}_s^1 \cap \mathcal{G}_s^2] \right) \\ &\leq \sum_{s=1}^S \left(\mathbb{P}[(\mathcal{G}_s^1 \cap \mathcal{G}_s^2)^c] + \mathbb{E}[\mathbb{P}[I_s = 0 \mid \mathcal{F}_{s-1/2}] \mathbf{1}_{\mathcal{G}_s^1 \cap \mathcal{G}_s^2}] \right) = o(n). \end{aligned} \quad (4.29)$$

Define the increment of the target quantity $\text{Overlap}(\pi)$ in the s -th step by

$$\mathcal{O}_s = \sum_{i < j, (i,j) \in \mathfrak{A}(\mathbf{M}_{s+1}, 2) \setminus \mathfrak{A}(\mathbf{M}_s, 2)} G_{i,j} \mathbf{G}_{\pi(i), \pi(j)}.$$

Then $\{I_s = 1\} \subset \{\mathcal{O}_s \geq \zeta\}$ for all $1 \leq i \leq S$. Therefore,

$$\begin{aligned} \text{Overlap}(\pi) &\geq \sum_{s=1}^S \mathcal{O}_s \geq \zeta \sum_{s=1}^S \mathbf{1}_{\{I_s=1\}} = \zeta S - \zeta \sum_{s=1}^S \mathbf{1}_{\{I_s=0\}} \\ &\geq \frac{1 - 2\eta - o(1)}{2\alpha_\eta - 1} n - \zeta \sum_{s=1}^S \mathbf{1}_{\{I_s=0\}}. \end{aligned}$$

Since $\frac{1-2\eta}{2\alpha_\eta-1} > \frac{1-\varepsilon}{2\alpha-1}$ by the choice of η in (4.7) and Lemma 4.6 (i), we conclude from Markov's inequality that

$$\mathbb{P} \left[\text{Overlap}(\pi) \leq \frac{1-\varepsilon}{2\alpha-1} n \right] \leq \frac{\zeta \cdot \mathbb{E} \left[\sum_{i=1}^S \mathbf{1}_{\{I_s=0\}} \right]}{\left(\frac{1-2\eta-o(1)}{2\alpha_\eta-1} - \frac{1-\varepsilon}{2\alpha-1} \right) n} \stackrel{(4.29)}{=} o(1),$$

completing the proof of the theorem. \square

4.3 Complimentary proofs

In this section, we prove Propositions 4.8, 4.9 and 4.10. We need some more definitions and notations. For a subgraph $\mathbf{H} \subset \mathbf{T}$, we define its capacity by

$$\text{Cap}(\mathbf{H}) = |V(\mathbf{H}) \cap \mathbf{Q}| - \alpha_\eta |E(\mathbf{H})|.$$

For a subtree $\mathbf{T}_0 \subset \mathbf{T}$, we say it is *dense* if $\text{Cap}(\mathbf{F}) > \text{Cap}(\mathbf{T})$ for any subgraph $\mathbf{F} \subsetneq \mathbf{T}_0$ with $V(\mathbf{T}_0) \cap \mathbf{L} \subset \mathbf{F}$. In addition, we choose a subgraph $\mathbf{F}_0 \subset \mathbf{T}_0$ with $V(\mathbf{T}_0) \cap \mathbf{L} \subset V(\mathbf{F}_0)$ such that $\text{Cap}(\mathbf{F}_0)$ is minimal (among all such choices for \mathbf{F}_0). We define a quantity $\mathbb{D}(\mathbf{T}_0)$ by

$$\log_n \mathbb{D}(\mathbf{T}_0) = \text{Cap}(\mathbf{T}_0) - \text{Cap}(\mathbf{F}_0). \quad (4.30)$$

Throughout this section, we will frequently deal with the conditional probability that \mathbf{G} contains a subgraph $\mathbf{T} \cong \mathbf{T}$ with fixed leaves given the existence of certain subgraph \mathbf{H} in \mathbf{G} . To this end, we decompose such event according to the intersecting pattern of $\mathbf{T} \cap \mathbf{H}$. For each possible realization \mathbf{F} of $\mathbf{T} \cap \mathbf{H}$, take $\mathbf{F} \subset \mathbf{T}$ such that \mathbf{F} is the image of \mathbf{F} under the isomorphism from \mathbf{T} to \mathbf{T} . Then it is straightforward from Markov's inequality that

$$\mathbb{P}[\exists \mathbf{T} \subset \mathbf{G} \text{ such that } \mathbf{T} \cong \mathbf{T} \text{ and } \mathbf{T} \cap \mathbf{H} = \mathbf{F} \mid \mathbf{H} \subset \mathbf{G}] \leq n^{|V(\mathbf{T}) \setminus V(\mathbf{F})|} p_\eta^{|E(\mathbf{T}) \setminus E(\mathbf{F})|}. \quad (4.31)$$

(In the above, \mathbf{H} is a fixed subgraph with vertex set in \mathbf{V} , and the event $\mathbf{H} \subset \mathbf{G}$ means that every edge in \mathbf{H} is contained in \mathbf{G} .) Thus, in order to upper-bound $\mathbb{P}[\exists \mathbf{T} \subset \mathbf{G} \text{ such that } \mathbf{T} \cong \mathbf{T} \mid \mathbf{H} \subset \mathbf{G}]$, we may sum over the right hand side of (4.31) over all possible \mathbf{F} .

Finally, we introduce some notations. For $1 \leq s \leq S$, fix a realization of \mathcal{F}_{s-1} . Recall the definitions of Fail_{s-1} , Suc_{s-1} , Suc_{s-1} and $\mathcal{A}_s^1, \mathcal{A}_s^2$ in (4.20), (4.21) and (4.22). Clearly \mathcal{A}_s^1 is decreasing and \mathcal{A}_s^2 is increasing. We let (below $L \cup Q$ means the set of all vertices appearing in L and Q)

$$O_{s-1} = \bigcup_{(L,Q) \in \text{Suc}_{s-1}} \{v : v \in L \cup Q\} \quad \text{and} \quad O_{s-1} = \bigcup_{(L,Q) \in \text{Suc}_{s-1}} \{v : v \in L \cup Q\}.$$

Then O_{s-1} is the set of vertices involved in the event \mathcal{A}_s^2 . From the algorithm we see

$$O_{s-1} \subset M_{s-1} = I_G(M_{s-1}), \quad O_{s-1} \subset \mathbf{M}_{s-1} = \Pi(M_{s-1}). \quad (4.32)$$

We define GO_{s-1} (respectively \mathbf{GO}_{s-1}) as the graph on M_{s-1} (respectively \mathbf{M}_{s-1}) which is the union of the trees that certify the events $\{L \bowtie_G Q\}$ for $(L, Q) \in \text{Suc}_{s-1}$ (respectively $\{\mathbf{L} \bowtie_{\mathbf{G}} \mathbf{Q}\}$ for $(\mathbf{L}, \mathbf{Q}) \in \text{Suc}_{s-1}$). Then it is clear that the event \mathcal{A}_s^2 is equivalent to $\mathbf{GO}_{s-1} \subset \mathbf{G}$. For a given realization of \mathcal{F}_{s-1} , we see that $O_{s-1}, \mathbf{O}_{s-1}, GO_{s-1}$ and \mathbf{GO}_{s-1} are deterministic, and in addition $GO_{s-1} \cong \mathbf{GO}_{s-1}$ through $\Pi \circ I_G^{-1}$ restricted on M_{s-1} . In addition, the rule of Algorithm 1 implies

$$\Delta(GO_s) \leq \kappa_0 + \Delta(\mathbf{T}), \quad (4.33)$$

where $\Delta(\cdot)$ is the maximal vertex degree. This is because except for the first time when a vertex is “used” it will always participate as a leaf vertex and thus only contributes 1 to the degree. For simplicity, we will denote $\widehat{\mathbb{P}}$ for the conditional probability on \mathbf{G} given by $\mathbb{P}[\cdot | \mathcal{A}_s^2] = \mathbb{P}[\cdot | \mathbf{GO}_s \subset \mathbf{G}]$.

4.3.1 Proof of Proposition 4.8

We continue to fix a realization of \mathcal{F}_{s-1} as above and also fix a ξ -tuple $\mathbf{L} \in \mathfrak{A}(\mathbf{M}_{s-1}, \xi)$. We first show the upper bound (4.24). Note that $\widehat{\mathbb{P}}$ is a product measure which admits the FKG inequality. As a result,

$$\begin{aligned} \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{A}_s^1, \mathcal{A}_s^2] &= \widehat{\mathbb{P}}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{A}_s^1] \\ &\leq \widehat{\mathbb{P}}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] = \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}], \end{aligned}$$

where the inequality holds because $\{\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}\}$ is increasing and \mathcal{A}_s^1 is decreasing, and the last equality follows from independence (recall (4.32)). This gives the upper bound.

Now we turn to the lower bound (4.25). The precise definitions of \mathcal{G}_s^1 and s -good tuples will be given later in this subsection, and we just assume \mathcal{F}_{s-1} satisfies \mathcal{G}_s^1 and \mathbf{L} is s -good for now. Let \mathcal{U} be the event that there is a unique χ -tuple $\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)$ such that $\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})$. Then we have

$$\begin{aligned} \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{A}_s^1, \mathcal{A}_s^2] &\geq \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathcal{U} | \mathcal{A}_s^1, \mathcal{A}_s^2] \\ &= \frac{1}{\widehat{\mathbb{P}}[\mathcal{A}_s^1]} \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \widehat{\mathbb{P}}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}, \mathcal{A}_s^1]. \end{aligned} \quad (4.34)$$

For each term in the preceding sum, note that \mathcal{U} is decreasing conditioned on $\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})$ and thus we have

$$\begin{aligned}\widehat{\mathbb{P}}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}, \mathcal{A}_s^1] &= \widehat{\mathbb{P}}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \cdot \widehat{\mathbb{P}}[\mathcal{A}_s^1 | \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \\ &= \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \cdot \widehat{\mathbb{P}}[\mathcal{A}_s^1 | \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \\ &\geq \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \cdot \widehat{\mathbb{P}}[\mathcal{A}_s^1 | \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})],\end{aligned}\quad (4.35)$$

where the second equality follows from independence and the last inequality follows from FKG inequality ($\widehat{\mathbb{P}}[\cdot | \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})]$ is still a product measure). Thus to prove the (4.25), the key is to show the following two lemmas.

Lemma 4.11. *For any $\mathbf{L} \in \mathfrak{A}(\mathbf{M}_{s-1}, \xi)$, it holds that*

$$\mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \geq [1 - o(1)] \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})]. \quad (4.36)$$

Lemma 4.12. *On some good event \mathcal{G}_s^1 (defined in Definition 4.14 below), we have that for any s -good $\mathbf{L} \in \mathfrak{A}(\mathbf{M}_{s-1}, \xi)$,*

$$\widehat{\mathbb{P}}[\mathcal{A}_s^1 | \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \geq [1 - o(1)] \widehat{\mathbb{P}}[\mathcal{A}_s^1]. \quad (4.37)$$

We now continue to complete the proof for (4.25). Note that

$$\mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] \leq \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})].$$

Combined with (4.34), (4.35) and Lemmas 4.11 and 4.12, this verifies (4.25).

We next prove (4.26). The upper bound follows easily from Markov's inequality. As for the lower bound, by Lemma 4.11 we have

$$\begin{aligned}\mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] &\geq \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathcal{U}] \\ &\geq [1 - o(1)] \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \gtrsim n^\chi p_\eta^\zeta.\end{aligned}\quad (4.38)$$

Combined with (4.24) and (4.25), it yields (4.26), completing the proof of Proposition 4.8.

It remains to prove Lemmas 4.11 and 4.12.

Proof of Lemma 4.11. (4.36) is equivalent to $\mathbb{P}[\mathcal{U} \mid \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \geq 1 - o(1)$. We now show $\mathbb{P}[\mathcal{U}^c \mid \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] = o(1)$ by a union bound. Let \mathbf{T} denote the tree in \mathbf{G} that certifies $\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})$ (so in particular $V(\mathbf{T}) = \{\mathbf{v} : \mathbf{v} \in \Pi(\mathbf{L}) \cup I_{\mathbf{G}}(\mathbf{Q})\}$ and $\mathbf{L} = \Pi(\mathbf{L})$ is the leaf set of \mathbf{T}). On the event \mathcal{U}^c , there exists another tree $\mathbf{T}' \cong \mathbf{T}$ with $V(\mathbf{T}) \neq V(\mathbf{T}')$ such that $\mathbf{T}' \subset \mathbf{G}$ and the leaf set of \mathbf{T}' is also \mathbf{L} . Therefore, \mathbf{T} and \mathbf{T}' intersect at some subgraph of \mathbf{F} with $\mathbf{L} \subset \mathbf{F} \subsetneq \mathbf{T}$. By a union bound, we see

$$\begin{aligned} & \mathbb{P}[\mathcal{U}^c \mid \Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \\ & \leq \sum_{\mathbf{L} \subset \mathbf{F} \subsetneq \mathbf{T}} \mathbb{P}[\text{there exists } \mathbf{T}' \subset \mathbf{G} \text{ such that } \mathbf{T}' \cong \mathbf{T} \text{ and } \mathbf{T}' \cap \mathbf{T} = \mathbf{F} \mid \mathbf{T} \subset \mathbf{G}] \\ & \stackrel{(4.31)}{\leq} \sum_{\mathbf{L} \subset \mathbf{F} \subsetneq \mathbf{T}} n^{|V(\mathbf{T}) \setminus V(\mathbf{F})|} p_{\eta}^{|E(\mathbf{T}) \setminus E(\mathbf{F})|} = \sum_{\mathbf{L} \subset \mathbf{F} \subsetneq \mathbf{T}} n^{|V(\mathbf{T}) \setminus V(\mathbf{F})| - \alpha_{\eta} |E(\mathbf{T}) \setminus E(\mathbf{F})|}, \end{aligned} \quad (4.39)$$

which is $o(1)$ by (4.9), as desired. \square

The rest of this subsection is devoted to the proof of Lemma 4.12. Recalling the definition of \mathbf{GO}_{s-1} , we see $\widehat{\mathbb{P}}$ is a product measure on the space $\Omega = \{0, 1\}^{E_0 \setminus E(\mathbf{GO}_{s-1})}$. Denote $\mathbf{T} \cong \mathbf{T}$ as the subtree of \mathbf{G} that certifies $\{\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})\}$ as before. For $\omega \in \Omega$, we write $\omega = \omega_1 \oplus \omega_2$, where $\omega_1 \in \Omega_1 = \{0, 1\}^{E_0 \setminus (E(\mathbf{GO}_{s-1}) \cup E(\mathbf{T}))}$ and $\omega_2 \in \Omega_2 = \{0, 1\}^{E(\mathbf{T})}$ (here \oplus means concatenation). It is then clear from the definition that

$$\widehat{\mathbb{P}}[\omega_1 \oplus \omega_2] = \widehat{\mathbb{P}}[\omega_1] \widehat{\mathbb{P}}[\omega_2] \text{ and } \{\Pi(\mathbf{L}) \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})\} = \{\omega_2 = (1, \dots, 1)\}.$$

For $i \in \{0, 1\}$, define

$$\mathcal{A}_s^i = \{\omega_1 \in \Omega_1 : \omega_1 \oplus \{i, \dots, i\} \in \mathcal{A}_s^1\}. \quad (4.40)$$

The fact that \mathcal{A}_s^1 is decreasing implies that \mathcal{A}_s^0 is also decreasing and $\mathcal{A}_s^1 \subset \mathcal{A}_s^0$. The left hand side of (4.37) can be expressed as $\widehat{\mathbb{P}}[\mathcal{A}_s^1]$. For $\widehat{\mathbb{P}}[\mathcal{A}_s^1]$ we have

$$\begin{aligned} \widehat{\mathbb{P}}[\mathcal{A}_s^1] &= \sum_{\omega_1 \in \Omega_1} \sum_{\omega_2 \in \Omega_2} \mathbf{1}_{\mathcal{A}_s^1}(\omega_1 \oplus \omega_2) \widehat{\mathbb{P}}[\omega_1 \oplus \omega_2] \\ &\leq \sum_{\omega_1 \in \Omega_1} \mathbf{1}_{\mathcal{A}_s^1}(\omega_1 \oplus \{0, \dots, 0\}) \widehat{\mathbb{P}}[\omega_1] \sum_{\omega_2 \in \Omega_2} \widehat{\mathbb{P}}[\omega_2] \\ &= \widehat{\mathbb{P}}[\{\omega_1 : \omega_1 \oplus \{0, \dots, 0\} \in \mathcal{A}_s^1\}] = \widehat{\mathbb{P}}[\mathcal{A}_s^0]. \end{aligned}$$

Provided with this, we see (4.37) reduces to

$$\frac{\widehat{\mathbb{P}}[\mathcal{A}_s^1]}{\widehat{\mathbb{P}}[\mathcal{A}_s^0]} = \widehat{\mathbb{P}}[\mathcal{A}_s^1 \mid \mathcal{A}_s^0] \geq 1 - o(1).$$

To this end, we note that for $\omega_1 \in \mathcal{A}_s^0 \setminus \mathcal{A}_s^1$, we have \mathcal{A}_s^1 holds if the edges in $E(\mathbf{T})$ are all closed while \mathcal{A}_s^1 fails after opening these edges. There are two possible scenarios for this: (i) opening edges in \mathbf{T} changes the realization of \mathcal{F}_{s-1} and hence alters the event \mathcal{A}_s^1 ; (ii) opening edges in \mathbf{T} does not change the realization of \mathcal{F}_{s-1} , but some of these edges participate in the tuple which certifies the failure of \mathcal{A}_s^1 . Denote by \mathcal{E}_s the event that after opening edges in $E(\mathbf{T})$, there exists a subgraph $\mathbf{T}' \subset \mathbf{G}$ with an isomorphism $\phi : \mathbf{T} \rightarrow \mathbf{T}'$ such that

$$E(\mathbf{T}') \cap E(\mathbf{T}) \neq \emptyset \text{ and } \phi(\mathbf{L}) = \Pi(\mathbf{L}') \text{ for some } \mathbf{L}' \in \bigcup_{1 \leq t \leq s-1} \text{EXP}_t. \quad (4.41)$$

Clearly, \mathcal{E}_s is an increasing event. We claim that both scenarios imply \mathcal{E}_s . Indeed, if \mathcal{E}_s does not hold, then after opening edges in $E(\mathbf{T})$, the realization of \mathcal{F}_{s-1} remains to be the same by the rule of Algorithm 1. So, in particular Fail_{s-1} remains to be the same. In addition, from the definition of \mathcal{A}_s^1 , we see under the event \mathcal{E}_s^c , whether \mathcal{A}_s^1 happens does not depend on openness/closedness for edges in $E(\mathbf{T})$. This proves the claim and now it suffice to show

$$\widehat{\mathbb{P}}[\mathcal{E}_s \mid \mathcal{A}_s^0] \leq \widehat{\mathbb{P}}[\mathcal{E}_s] = o(1),$$

where the first inequality follows from FKG.

We divide \mathcal{E}_s according to the intersecting patterns of \mathbf{T}' with \mathbf{T} and GO_{s-1} . Denote \mathcal{P} for the pairs $(\mathbf{F}_1, \mathbf{F}_2)$ with $\mathbf{F}_1, \mathbf{F}_2 \subset \mathbf{T}$ such that \mathbf{F}_1 is a *proper* subtree of \mathbf{T} with at least one edge, $V(\mathbf{F}_1) \cap V(\mathbf{F}_2) = \emptyset$ and $\mathbf{L} \subset V(\mathbf{F}_1) \cup V(\mathbf{F}_2)$. For a pair $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}$, define $\mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2)$ be the event that (4.41) holds and that

$$\phi(E(\mathbf{F}_1)) \cap E(\mathbf{T}) \neq \emptyset, \quad \phi(\mathbf{F}_1 \cup \mathbf{F}_2) = \mathbf{T}' \cap (\text{GO}_{s-1} \cup \mathbf{T}).$$

Recalling (4.41), we obtain

$$\mathcal{E}_s \subset \mathcal{E}_s(\mathbf{T}, \emptyset) \cup \bigcup_{(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}} \mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2). \quad (4.42)$$

(Indeed, \mathbf{F}_1 can be taken as any component of $\phi^{-1}(\mathbf{T}' \cap (\mathbf{T} \cup \text{GO}_{s-1}))$ which contains at least one edge in $E(\mathbf{T})$.) We want to exclude the case $\mathcal{E}_s(\mathbf{T}, \emptyset)$ from the assumption that \mathbf{L} is s -good. To this end, we make the following definition.

Definition 4.13. For a ξ -tuple $L_0 \in \mathfrak{A}(M_{s-1}, \xi)$, we say it is s -bad, if for some $T \cong \mathbf{T}$ with leaf set $L_0 = I_G(L_0)$ and $V(T) \cap M_{s-1} = L_0$, we have that the graph $T \cup \text{GO}_{s-1}$ contains a subgraph

$T^* \cong \mathbf{T}$ satisfying that $E(T^*) \cap E(T) \neq \emptyset$ and that the leaf set of T^* equals to $L^* = I_G(L^*)$ for some $L^* \in \bigcup_{1 \leq t \leq s-1} \text{EXP}_t$ (note that in this definition for ‘some’ $T \cong \mathbf{T}$ is equivalent to for ‘any’ $T \cong \mathbf{T}$). Otherwise we say $L_0 \in \mathfrak{A}(M_{s-1}, \xi)$ is s -good.

From the definition and the fact that $GO_{s-1} \cong \mathbf{GO}_{s-1}$ through $\Pi \circ I_G^{-1}$, we see that $\mathcal{E}_s \cap \mathcal{E}_s(\mathbf{T}, \emptyset) = \emptyset$ under the assumption that L is s -good. Thus, (4.42) can be strengthened as

$$\mathcal{E}_s \subset \bigcup_{(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}} \mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2). \quad (4.43)$$

For a pair $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}$, let $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$ be the number of ξ -tuples $L' \in \bigcup_{1 \leq t \leq s-1} \text{EXP}_t$ such that there exist two embeddings $\phi_i : \mathbf{F}_i \rightarrow \mathbf{T} \cup \mathbf{GO}_{s-1}$, $i = 1, 2$ such that

$$\phi_1(E(\mathbf{F}_1)) \cap E(\mathbf{T}) \neq \emptyset \text{ and } \phi_1(V(\mathbf{F}_1) \cap \mathbf{L}) \cup \phi_2(V(\mathbf{F}_2) \cap \mathbf{L}) = \Pi(L').$$

(Note that $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$ depends on \mathbf{T} although we did not include \mathbf{T} in the notation.) Then $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$ is the number of possible choices for the leaf set of \mathbf{T}' which may potentially certify the event $\mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2)$. For each fixed choice, we see from (4.31) that the probability that this indeed certifies the event $\mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2)$ is upper-bounded by

$$\text{Prob}(\mathbf{F}_1, \mathbf{F}_2) = n^{|V(\mathbf{T})| - |V(\mathbf{F}_1)| - |V(\mathbf{F}_2)|} p_\eta^{|E(\mathbf{T})| - |E(\mathbf{F}_1)| - |E(\mathbf{F}_2)|}. \quad (4.44)$$

Then by a union bound, for any pair $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}$ it holds that

$$\widehat{\mathbb{P}}[\mathcal{E}_s(\mathbf{F}_1, \mathbf{F}_2)] \leq \text{Enum}(\mathbf{F}_1, \mathbf{F}_2) \times \text{Prob}(\mathbf{F}_1, \mathbf{F}_2). \quad (4.45)$$

Let \mathcal{P}_0 be the collection of $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}$ which maximizes the right hand side of (4.45). We claim that for any $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}_0$ the components of \mathbf{F}_2 are dense trees which intersect \mathbf{L} . We now prove this claim by contradiction, and we divide the proof into two cases. If \mathbf{F}_2 contains a tree component disjoint from \mathbf{L} , we can simply remove this component from \mathbf{F}_2 to get a pair $(\mathbf{F}_1, \mathbf{F}'_2) \in \mathcal{P}$ with

$$\text{Prob}(\mathbf{F}_1, \mathbf{F}'_2) > \text{Prob}(\mathbf{F}_1, \mathbf{F}_2), \quad \text{Enum}(\mathbf{F}_1, \mathbf{F}'_2) \geq \text{Enum}(\mathbf{F}_1, \mathbf{F}_2),$$

contradicting the maximality. If \mathbf{F}_2 contains a tree component \mathbf{T}_0 which intersects \mathbf{L} but is not dense, we may find some $\mathbf{F}_0 \subsetneq \mathbf{T}_0$ which contains $V(\mathbf{T}_0) \cap \mathbf{L}$ such that $\text{Cap}(\mathbf{F}_0) < \text{Cap}(\mathbf{T}_0)$ (this is feasible since α_η is irrational). We define \mathbf{F}'_2 by replacing \mathbf{T}_0 with \mathbf{F}_0 in \mathbf{F}_2 and get a pair $(\mathbf{F}_1, \mathbf{F}'_2) \in \mathcal{P}$. Again, it satisfies

$$\text{Prob}(\mathbf{F}_1, \mathbf{F}'_2) > \text{Prob}(\mathbf{F}_1, \mathbf{F}_2) \text{ and } \text{Enum}(\mathbf{F}_1, \mathbf{F}'_2) \geq \text{Enum}(\mathbf{F}_1, \mathbf{F}_2),$$

contradicting the maximality. This completes the verification of the claim. Recall (4.30). We are now ready to define our good event \mathcal{G}_s^1 .

Definition 4.14. Fix some large constant $\kappa_1 = \kappa_1(\eta, \alpha_\eta, \mathbf{T})$ which will be determined later. We define \mathcal{G}_s^1 to be the event that for any $(\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}_0$ and for any $\mathbf{T} \subset \mathbf{G}$,

$$\text{Enum}(\mathbf{F}_1, \mathbf{F}_2) \leq \begin{cases} \kappa_1 n \times (np_\eta)^{|E(\mathbf{T})| - |E(\mathbf{F}_2)|}, & \text{if } V(\mathbf{F}_1) \cap \mathbf{L} = \emptyset, \\ \kappa_1 \mathbb{D}(\mathbf{F}_1) \times (np_\eta)^{|E(\mathbf{T})| - |E(\mathbf{F}_1)| - |E(\mathbf{F}_2)|}, & \text{if } V(\mathbf{F}_1) \cap \mathbf{L} \neq \emptyset. \end{cases} \quad (4.46)$$

Since $|\mathcal{P}|$ is uniformly bounded in n , it suffices to show that on \mathcal{G}_s^1

$$\text{Enum}(\mathbf{F}_1, \mathbf{F}_2) \times \text{Prob}(\mathbf{F}_1, \mathbf{F}_2) = o(1) \text{ for any } (\mathbf{F}_1, \mathbf{F}_2) \in \mathcal{P}_0. \quad (4.47)$$

To this end, we may write $(\mathbf{F}_1, \mathbf{F}_2) = (\mathbf{T}_0, \mathbf{T}_1 \cup \dots \cup \mathbf{T}_l)$, where \mathbf{T}_i is a subtree of \mathbf{T} for $0 \leq i \leq l$. The proof is divided into two cases.

In the case that $V(\mathbf{T}_0) \cap \mathbf{L} = \emptyset$, the target product is upper-bounded by (recalling (4.46))

$$\begin{aligned} & \kappa_1 n^{\chi + \zeta} p_\eta^{2\zeta} \times n^{-|V(\mathbf{T}_0)| + 1} p_\eta^{-|E(\mathbf{T}_0)|} \times \prod_{i=1}^l n^{-|V(\mathbf{T}_i) \cap \mathbf{Q}| - |E(\mathbf{T}_i)|} p_\eta^{-2|E(\mathbf{T}_i)|} \\ &= \kappa_1 n^{\chi - (2\alpha_\eta - 1)\zeta} \times (np_\eta)^{-|E(\mathbf{T}_0)|} \times \prod_{i=1}^l n^{-|V(\mathbf{T}_i) \cap \mathbf{Q}| + (2\alpha_\eta - 1)|E(\mathbf{T}_i)|}. \end{aligned}$$

Note that the second term (i.e., $(np_\eta)^{-|E(\mathbf{T}_0)|}$) is no more than $n^{\alpha_\eta - 1}$ since $E(\mathbf{T}_0) \neq \emptyset$ and that the third term is bounded by 1 from (4.10). Thus, we see the expression above is upper-bounded by $\kappa_1 n^{\chi - (2\alpha_\eta - 1)\zeta} \times n^{\alpha_\eta - 1}$, which is $o(1)$ by Lemma 4.6 (i). This proves the desired bound in this case.

In the case that $V(\mathbf{T}_0) \cap \mathbf{L} \neq \emptyset$, we may write

$$\mathbb{D}(\mathbf{T}_0) = n^{\text{Cap}(\mathbf{T}_0) - \text{Cap}(\mathbf{F}_0)} = n^{\text{Cap}(\mathbf{T}_0) - \sum_{j=1}^m \text{Cap}(\mathbf{T}'_j)},$$

where $\mathbf{F}_0 \subset \mathbf{T}_0$ is the union of disjoint trees $\mathbf{T}'_1, \dots, \mathbf{T}'_m$ with $V(\mathbf{T}_0) \cap \mathbf{L} \subset V(\mathbf{F}_0)$ such that $\text{Cap}(\mathbf{F}_0)$ is minimized. As a result, we see the product $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2) \times \text{Prob}(\mathbf{F}_1, \mathbf{F}_2)$ is upper-bounded by

$$\kappa_1 n^{\chi + \zeta} p_\eta^{2\zeta} \times n^{-\sum_{j=1}^m \text{Cap}(\mathbf{T}'_j)} \times (np_\eta)^{-|E(\mathbf{T}_0)|} \times \prod_{i=1}^l n^{-|V(\mathbf{T}_i) \cap \mathbf{Q}| - |E(\mathbf{T}_i)|} p_\eta^{-2|E(\mathbf{T}_i)|}. \quad (4.48)$$

If some of \mathbf{T}'_j is not a singleton, then (4.48) is upper-bounded by (recalling that the product above is upper-bounded by 1)

$$\begin{aligned} & \kappa_1 n^{\chi+\zeta} p_\eta^{2\zeta} \times n^{-\sum_{j=1}^m \text{Cap}(\mathbf{T}'_j)} \times (np_\eta)^{-\sum_{j=1}^m |E(\mathbf{T}'_j)|} \\ & \leq \kappa_1 n^{\chi-(2\alpha_\eta-1)\zeta} \prod_{j=1}^m n^{-|V(\mathbf{T}'_j) \cap Q(\mathbf{T})| + (2\alpha_\eta-1)|E(\mathbf{T}'_j)|}, \end{aligned}$$

which is $o(1)$ by Lemma 4.6 (i) and (iv). If each \mathbf{T}'_j is a singleton, then (4.48) is upper-bounded by

$$\kappa_1 n^{\chi+\zeta} p_\eta^{2\zeta} \times (np_\eta)^{-1} \leq \kappa_1 n^{\chi-(2\alpha_\eta-1)\zeta} \times n^{\alpha_\eta-1},$$

which is also $o(1)$ Lemma 4.6 (i). This completes the proof in this case, and thus finally completes the proof of Lemma 4.12.

4.3.2 Proof of Proposition 4.9

We continue to fix some $1 \leq s \leq S$ and a realization of \mathcal{F}_{s-1} . We further fix a realization of CAND_s , and hence we also get the set of s -good ξ -tuples in CAND_s , denoted as $\text{GC}_s = \{\mathbf{L}_1, \dots, \mathbf{L}_l\}$. For any nonempty subset $\mathbf{R} \subset \mathbf{L}$, denote $\text{Span}(\mathbf{R})$ for the subtree of \mathbf{T} spanned by \mathbf{R} (i.e., the minimal subtree that contains \mathbf{R}). Throughout this section, it will be convenient to partition pairs in $\text{GC}_s \times \text{GC}_s$ according to their intersecting patterns. More precisely, for two ξ -tuples $\mathbf{L} = (t_1, \dots, t_\xi)$ and $\mathbf{L}' = (t'_1, \dots, t'_\xi)$, we let $\text{Loc}(\mathbf{L}, \mathbf{L}') = \{1 \leq i \leq \xi : t'_i \in \{t_1, \dots, t_\xi\}\}$. For any subset $\mathbf{R} \subset \mathbf{L}$ (Recall that $\mathbf{L} = \{\chi+1, \dots, \chi+\xi\}$), we let

$$\text{IP}_s(\mathbf{R}) = \{(\mathbf{L}_i, \mathbf{L}_j) \in \text{GC}_s \times \text{GC}_s : \text{Loc}(\mathbf{L}_i, \mathbf{L}_j) = \{r - \chi, r \in \mathbf{R}\}\}.$$

We are now ready to define our good event \mathcal{G}_s^2 .

Definition 4.15. Fix some positive constants κ_2, κ_3 depending only on η, α_η and \mathbf{T} which will be determined later. The good event \mathcal{G}_s^2 is the intersection of the following two events:

- (i) $l = |\text{GC}_s| \geq \kappa_2 (np_\eta)^\zeta$.
- (ii) For any nonempty subset $\mathbf{R} \subset \mathbf{L}$,

$$|\text{IP}_s(\mathbf{R})| \leq \kappa_3 (np_\eta)^\zeta \times \mathbb{D}(\text{Span}(\mathbf{R})) \times (np_\eta)^{|E(\mathbf{T}) \setminus E(\text{Span}(\mathbf{R}))|}.$$

We next assume that the realization $\mathcal{F}_{s-1/2} = \sigma(\mathcal{F}_s \cup \text{CAND}_s)$ satisfies $\mathcal{G}_s^1 \cap \mathcal{G}_s^2$ and prove (4.27). For simplicity, we write $\mathbf{L}_i = \Pi(\mathbf{L}_i)$. Since \mathcal{F}_{s-1} satisfies \mathcal{G}_s^1 , from Proposition 4.8 we see for

any $\mathbf{L}_i \in \Pi(\text{GC}_s)$,

$$\mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1/2}] = \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \geq [1 - o(1)] \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] \geq c_1 n^\chi p_\eta^\zeta.$$

Combined with (i) in \mathcal{G}_s^2 , this yields a lower bound on the right hand side of (4.27):

$$(\mathbb{E}[X_s \mid \mathcal{F}_{s-1/2}])^2 = \left(\sum_{i=1}^l \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1/2}] \right)^2 \gtrsim l^2 (n^\chi p_\eta^\zeta)^2 \gtrsim (n^{\chi+\zeta} p_\eta^{2\zeta})^2. \quad (4.49)$$

For the left hand side of (4.27), we may expand it out and break the sum into several parts as follows:

$$\begin{aligned} \mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1/2}] &= \sum_{i,j=1}^l \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1/2}] \\ &= \sum_{(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\emptyset)} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \end{aligned} \quad (4.50)$$

$$+ \sum_{\emptyset \neq \mathbf{R} \subset \mathbf{L}} \sum_{(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\mathbf{R})} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}]. \quad (4.51)$$

We estimate the probabilities in the sum above by the following two lemmas.

Lemma 4.16. *For any $(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\emptyset)$, it holds that*

$$\begin{aligned} &\mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \\ &\leq [1 + o(1)] \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}]. \end{aligned} \quad (4.52)$$

Lemma 4.17. *For any nonempty subset $\mathbf{R} \subset \mathbf{L}$, let $\mathbf{F}_{\mathbf{R}}$ be the subgraph of \mathbf{T} with $V(\mathbf{F}_{\mathbf{R}}) \cap \mathbf{L} = \mathbf{R}$ such that $\text{Cap}(\mathbf{F}_{\mathbf{R}})$ is minimized out of all such subgraphs. Then for any $(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\mathbf{R})$, it holds that*

$$\mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \lesssim (n^\chi p_\eta^\zeta)^2 \times n^{-\text{Cap}(\mathbf{F}_{\mathbf{R}})}. \quad (4.53)$$

Proof of Proposition 4.9. From Lemma 4.16, (4.50) is upper-bounded by

$$[1 + o(1)] \sum_{(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\emptyset)} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \quad (4.54)$$

$$\leq [1 + o(1)] \sum_{i,j=1}^l \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}] \quad (4.55)$$

which is $[1 + o(1)](\mathbb{E}(X_s | \mathcal{F}_{s-1}))^2$. In addition, for each nonempty subset $\mathbf{R} \subset \mathbf{L}$, by (ii) in \mathcal{G}_s^2 and Lemma 4.17 we see that $\sum_{(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\mathbf{R})} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{F}_{s-1}]$ is bounded by a constant multiplies

$$\begin{aligned} & (np_\eta)^\zeta \times \mathbb{D}(\text{Span}(\mathbf{R})) \times (np_\eta)^{|E(\mathbf{T}) \setminus E(\text{Span}(\mathbf{R}))|} \times (n^\chi p_\eta^\zeta)^2 \times n^{-\text{Cap}(\mathbf{F}_{\mathbf{R}})} \\ &= (n^{\chi+\zeta} p_\eta^{2\zeta})^2 \times n^{-|E(\text{Span}(\mathbf{R}))| + |V(\text{Span}(\mathbf{R})) \cap \mathbf{Q}|} \times n^{-2 \text{Cap}(\mathbf{F}_{\mathbf{R}})}, \end{aligned} \quad (4.56)$$

where we used the fact that $\mathbb{D}(\text{Span}(\mathbf{R})) = n^{\text{Cap}(\text{Span}(\mathbf{R})) - \text{Cap}(\mathbf{F}_{\mathbf{R}})}$. Suppose $\mathbf{F}_{\mathbf{R}}$ is a union of subtrees $\mathbf{T}_1, \dots, \mathbf{T}_r$ of \mathbf{T} . We note that

$$|E(\text{Span}(\mathbf{R}))| - |V(\text{Span}(\mathbf{R})) \cap \mathbf{Q}| = |\mathbf{R}| - 1 \geq |\mathbf{R}| - r = \sum_{i=1}^r (|E(\mathbf{T}_i)| - |V(\mathbf{T}_i) \cap \mathbf{Q}|),$$

Thus, the term $n^{-|E(\text{Span}(\mathbf{R}))| + |V(\text{Span}(\mathbf{R})) \cap \mathbf{Q}|} \times n^{-2 \text{Cap}(\mathbf{F}_{\mathbf{R}})}$ in (4.56) is bounded by

$$\prod_{i=1}^r n^{-|E(\mathbf{T}_i)| + |V(\mathbf{T}_i) \cap \mathbf{Q}| - 2 \text{Cap}(\mathbf{T}_i)} = \prod_{i=1}^r n^{-|V(\mathbf{T}_i) \cap \mathbf{Q}| + (2\alpha_\eta - 1)|E(\mathbf{T}_i)|},$$

which is $o(1)$ from Lemma 4.6 (iv). This shows that (4.56) is $o((n^{\chi+\zeta} p_\eta^{2\zeta})^2)$ for each \mathbf{R} . Since the number of possible \mathbf{R} is uniformly bounded in n , we complete the proof by combining with (4.49). \square

It remains to prove Lemma 4.16 and Lemma 4.17. We note that for any two tuples $\mathbf{L}_i, \mathbf{L}_j \in \Pi(\text{GC}_s)$, it holds

$$\begin{aligned} & \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{F}_{s-1}] = \widehat{\mathbb{P}}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{A}_s^1] \\ & \leq \widehat{\mathbb{P}}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] = \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}], \end{aligned} \quad (4.57)$$

where the inequality follows from the FKG inequality and the last equality follows from independence. Then it is clear that $\mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}, \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}]$ is upper-bounded by

$$\begin{aligned} & \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q}), \mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}] \\ &= \sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})]. \end{aligned} \quad (4.58)$$

By Proposition 4.8 and (4.38), we have

$$\sum_{\mathbf{Q} \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(\mathbf{Q})] \leq [1 + o(1)] \mathbb{P}[\mathbf{L}_i \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} | \mathcal{F}_{s-1}] \lesssim n^\chi p_\eta^\zeta. \quad (4.59)$$

Thus it remains to show for each $Q \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)$, we have

$$\mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)] \leq \begin{cases} [1 + o(1)] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1}], & \text{if } (\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\emptyset); \\ O(n^\chi p_\eta^\zeta \times n^{-\text{Cap}(\mathbf{F}_{\mathbf{R}})}), & \text{if } (\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\mathbf{R}), \mathbf{R} \neq \emptyset. \end{cases}$$

Fix such a tuple Q and denote $\mathbf{T} \subset \mathbf{G}$ the subtree that certifies $\{\mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)\}$. Similarly for each $Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)$, we let $\mathbf{T}' \subset \mathbf{G}$ be the subtree that certifies $\{\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q')\}$. Then from a union bound we get

$$\begin{aligned} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)] &\leq \sum_{Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi)} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q') \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)] \\ &= \sum_{\substack{Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi) \\ E(\mathbf{T}) \cap E(\mathbf{T}') = \emptyset}} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q')] + \sum_{\substack{Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi) \\ E(\mathbf{T}) \cap E(\mathbf{T}') \neq \emptyset}} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q') \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)], \end{aligned} \quad (4.60)$$

where the equality follows from independence.

For any subset $\mathbf{R} \subset \mathbf{L}$ (possibly $\mathbf{R} = \emptyset$), we define $\mathcal{P}(\mathbf{R})$ as the collection of nonempty subgraphs $\mathbf{F} \subset \mathbf{T}$ with $V(\mathbf{F}) \cap \mathbf{L} = \mathbf{R}$.

Proof of Lemma 4.16. For the case $(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\emptyset)$, we have $\mathbf{L}_i \cap \mathbf{L}_j = \emptyset$. The first sum in (4.60) is bounded by $[1 + o(1)] \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}]$ for the same reason as (4.59). As for the second sum in (4.60), it can be upper-bounded by

$$\sum_{\mathbf{F} \in \mathcal{P}(\emptyset)} \sum_{\substack{Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi) \\ \mathbf{T} \cap \mathbf{T}' \cong \mathbf{F}}} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q') \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)].$$

For each $\mathbf{F} \in \mathcal{P}(\emptyset)$, it is readily to see the second summation above is bounded by

$$n^{|Q \setminus V(\mathbf{F})|} p_\eta^{|E(\mathbf{T}) \setminus E(\mathbf{F})|} = n^\chi p_\eta^\zeta \times n^{-|V(\mathbf{F})|} p_\eta^{-|E(\mathbf{F})|} = o(n^\chi p_\eta^\zeta) = o(\mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} \mathbf{R}_{s-1} \mid \mathcal{F}_{s-1}]),$$

where we used the fact that $|V(\mathbf{F})| > |E(\mathbf{F})|$ for any $\mathbf{F} \in \mathcal{P}(\emptyset)$ since \mathbf{F} is a forest. Since the cardinality of $\mathcal{P}(\emptyset)$ is uniformly bounded in n , this concludes Lemma 4.16. \square

Proof of Lemma 4.17. For the case $(\mathbf{L}_i, \mathbf{L}_j) \in \text{IP}_s(\mathbf{R})$ with $\mathbf{R} \neq \emptyset$, similarly, the first sum in (4.60) remains to be $O(n^\chi p_\eta^\zeta)$. Note that the second sum can be expressed as

$$\sum_{\mathbf{F} \in \mathcal{P}(\mathbf{R})} \sum_{\substack{Q' \in \mathfrak{A}(\mathbf{R}'_{s-1}, \chi) \\ \mathbf{T} \cap \mathbf{T}' \cong \mathbf{F}}} \mathbb{P}[\mathbf{L}_j \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q') \mid \mathbf{L}_i \bowtie_{\mathbf{G}} I_{\mathbf{G}}(Q)] \leq O(n^\chi p_\eta^\zeta \times \max_{\mathbf{F} \in \mathcal{P}(\mathbf{R})} n^{-\text{Cap}(\mathbf{F})}).$$

Since $\text{Cap}(\mathbf{F}_{\mathbf{R}}) = \min_{\mathbf{F} \in \mathcal{P}(\mathbf{R})} \text{Cap}(\mathbf{F})$, this yields Lemma 4.17. \square

4.3.3 Proof of Proposition 4.10

We start with some notations. For any simple graph \mathbf{H} and $\mathbf{R} \subset V(\mathbf{H})$ with $|\mathbf{R}| = r \geq 1$, and any r -tuple $\mathbf{I} \in \mathfrak{A}([n], r)$, we define $\text{EXT}(\mathbf{I}; \mathbf{R}, \mathbf{H})$ to be the collection of subgraphs H in G satisfying the following condition: there is an isomorphism $\phi : \mathbf{H} \rightarrow H$ such that $\phi(\mathbf{R}) = (v_i)_{i \in \mathbf{I}}$ (note that here the equality is equal in the sense of a tuple, not just in the sense of a set). In addition, we let $\text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H}) = \text{EXT}(\mathbf{I}; \mathbf{R}, \mathbf{H})$. For a pair (\mathbf{R}, \mathbf{H}) , we say it is a *dense pattern*, if for any subgraph $\mathbf{H}' \subsetneq \mathbf{H}$ with $\mathbf{R} \subset V(\mathbf{H}')$,

$$|V(\mathbf{H}) \setminus V(\mathbf{H}')| < \alpha_\eta |E(\mathbf{H}) \setminus E(\mathbf{H}')|. \quad (4.61)$$

We say (\mathbf{R}, \mathbf{H}) is a *sparse pattern*, if for any subgraph $\mathbf{H}' \subset \mathbf{H}$ with $\mathbf{R} \subset V(\mathbf{H}')$ and $E(\mathbf{H}') \neq \emptyset$,

$$|V(\mathbf{H}') \setminus \mathbf{R}| > \alpha_\eta |E(\mathbf{H}')|. \quad (4.62)$$

Note that dense and sparse patterns are mutually exclusive (this can be checked by proof of contradiction with taking $\mathbf{H}' = \mathbf{R}$ in (4.61) and taking $\mathbf{H}' = \mathbf{H}$ in (4.62)) but they are not mutually complementary, and in addition sparse pattern corresponds to the α_η -safe extension in [40]. Recall (4.30) for the definition of $\mathbb{D}(\mathbf{T}_0)$ for $\mathbf{T}_0 \subset \mathbf{T}$. We introduce yet another good event \mathcal{G} as follow:

Definition 4.18. Fix a large constant $\kappa_4 = \kappa_4(\alpha_\eta, \mathbf{T})$ which will be determined later. Define \mathcal{G} as the event that for any $\emptyset \neq \mathbf{R} \subset \mathbf{L}$ and any tree $\mathbf{T}_0 \subset \mathbf{T}$ with $V(\mathbf{T}_0) \cap \mathbf{L} = \mathbf{R}$,

$$\max_{\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)} \text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{T}_0) \leq \kappa_4 \mathbb{D}(\mathbf{T}_0). \quad (4.63)$$

Lemma 4.19. For some large constant κ_4 , it holds that $\mathbb{P}[\mathcal{G}^c] = o(1)$.

Proof. Since the number of pairs $(\mathbf{R}, \mathbf{T}_0)$ is bounded in n , we only need to prove (4.63) holds with probability tending to 1 for any fixed pair. For a fixed pair $(\mathbf{R}, \mathbf{T}_0)$, we take a subgraph $\mathbf{F}_0 \subset \mathbf{T}_0$ with $V(\mathbf{F}_0) \cap \mathbf{L} = \mathbf{R}$ such that $\text{Cap}(\mathbf{F}_0)$ is minimized. Note that subgraphs in $\text{EXT}(\mathbf{I}; \mathbf{R}, \mathbf{T}_0)$ can be “constructed” as follows:

Step 1. choose a subgraph $F \cong \mathbf{F}_0$ in G with fixed leaves in \mathbf{R} ;

Step 2. add vertices and edges to F and get a final subgraph $T \cong \mathbf{T}$ in G .

Since both \mathbf{F}_0 and $\mathbf{T}_0 \setminus_\star \mathbf{F}_0$ are forests, we may assume $\mathbf{T}_1, \dots, \mathbf{T}_l$ are the components of \mathbf{F}_0 , and $\mathbf{T}'_1, \dots, \mathbf{T}'_m$ are the components of $\mathbf{T}_0 \setminus_\star \mathbf{F}_0$. Denote $\mathbf{R}_i = V(\mathbf{T}_i) \cap \mathbf{R}$ for $1 \leq i \leq l$ and $\mathbf{R}'_j = V(\mathbf{F}_0) \cap V(\mathbf{T}'_j)$ for $1 \leq j \leq m$. By the minimality of $\text{Cap}(\mathbf{F}_0)$ and the irrationality of α_η ,

we see that $(\mathbf{R}_i, \mathbf{T}_i)$ is a dense pattern for each $1 \leq i \leq l$ and $(\mathbf{R}'_j, \mathbf{T}'_j)$ is a sparse pattern for each $1 \leq j \leq m$. It then suffices to show the following two items.

- For any fixed dense pattern (\mathbf{R}, \mathbf{H}) , there exists a large constant κ such that probability $1 - o(1)$,

$$\max_{\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)} \text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H}) \leq \kappa. \quad (4.64)$$

- For any fixed sparse pattern (\mathbf{R}, \mathbf{H}) , there exists a large constant κ such that with probability $1 - o(1)$,

$$\max_{\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)} \text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H}) \leq \kappa n^{|V(\mathbf{H}) \setminus \mathbf{R}| - \alpha_\eta |E(\mathbf{H})|}. \quad (4.65)$$

By [40, Corollary 4] (where the condition α_η -safe is equivalent to the sparsity of (\mathbf{R}, \mathbf{H})), we see that (4.65) holds. Thus, it remains to prove (4.64). To this end, we will use the following intuition: for a dense pattern (\mathbf{R}, \mathbf{H}) , if $\text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H})$ is large for some $\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)$, then there exists a subgraph $K \subset G$ with bounded size such that $|V(K)| - \alpha_\eta |E(K)| < -1$. We next elaborate this precisely. Since (\mathbf{R}, \mathbf{H}) is dense, we have

$$-\delta \stackrel{\text{def}}{=} \max_{\mathbf{H}' \subsetneq \mathbf{H}, \mathbf{R} \subset V(\mathbf{H}')} (|V(\mathbf{H}) \setminus V(\mathbf{H}')| - \alpha_\eta |E(\mathbf{H}) \setminus E(\mathbf{H}')|) < 0. \quad (4.66)$$

If $\text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H})$ exceeds a large integer κ for some $\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)$, then there exist κ subgraphs $H_1, \dots, H_\kappa \cong \mathbf{H}$ in G , such that each isomorphism from \mathbf{H} to H_i maps \mathbf{R} to $(v_i)_{i \in \mathbf{I}}$. Let $K_i = H_1 \cup H_2 \cup \dots \cup H_i$ for $1 \leq i \leq \kappa$. We note that whenever $K_{i+1} \setminus K_i \neq \emptyset$, each of its component is isomorphic to some $\mathbf{H} \setminus_\star \mathbf{H}'$ with $\mathbf{H}' \subset \mathbf{H}$ and $\mathbf{R} \subset V(\mathbf{H}')$. Denoting $P(K_i) = |V(K_i)| - \alpha_\eta |E(K_i)|$, we then deduce from (4.66) that

- $P(K_{i+1}) \leq P(K_i)$ for any $i \geq 1$, with equality holds if and only if $K_{i+1} = K_i$,
- If $K_{i+1} \neq K_i$, then $P(K_{i+1}) \leq P(K_i) - \delta$.

In addition, for each fixed K_i , there exists a large integer $N = N(K_i)$ depending only on the size of K_i such that $K_{i+N} \neq K_i$. Combined with the fact that $P(K_1) = |V(H)| - \alpha_\eta |E(H)|$ is bounded in n , we see $P(K_\kappa) < -1$ for κ large enough. Clearly, we also have $|V(K_\kappa)| \leq \kappa |V(\mathbf{H})|$. Therefore,

$$\begin{aligned} & \mathbb{P} \left[\max_{\mathbf{I} \in \mathfrak{A}([n], |\mathbf{R}|)} \text{Ext}(\mathbf{I}; \mathbf{R}, \mathbf{H}) \geq \kappa \right] \\ & \leq \mathbb{P} [\exists K \subset G \text{ with } |V(K)| \leq \kappa |V(\mathbf{H})| \text{ and } P(K) < -1] \\ & \leq \sum_{\substack{|V(\mathbf{K})| \leq \kappa |V(\mathbf{H})| \\ P(\mathbf{K}) < -1}} n^{|V(\mathbf{K})|} p_\eta^{|E(\mathbf{K})|} = \sum_{\substack{|V(\mathbf{K})| \leq \kappa |V(\mathbf{H})| \\ P(\mathbf{K}) < -1}} n^{P(\mathbf{K})} = O(n^{-1}). \end{aligned} \quad (4.67)$$

This proves (4.64), and thus completes the proof of the lemma. \square

We will prove Proposition 4.10 by induction on s . Assuming for some fixed $1 \leq s \leq S$ it holds uniformly that for all $1 \leq t \leq s-1$,

$$\mathbb{P}[(\mathcal{G}_t^1)^c] + \mathbb{P}[(\mathcal{G}_t^2)^c] = o(1), \quad (4.68)$$

we will show that (4.68) holds for $t = s$ (This will in particular prove the base case of $t = 1$). Recall definition (4.23) for X_t , $1 \leq t \leq s-1$, the induction hypothesis (4.68) is used in the following lemma:

Lemma 4.20. *Assuming (4.68) for $1 \leq t \leq s-1$ it holds that*

$$\mathbb{P}[X_t < \log n] = o(1) \text{ uniformly for all } 1 \leq t \leq s-1. \quad (4.69)$$

Proof. From the induction hypothesis we see $\mathbb{P}[X_t < \log n]$ is no more than

$$\mathbb{P}[(\mathcal{G}_t^1)^c] + \mathbb{P}[(\mathcal{G}_t^2)^c] + \mathbb{P}[X_t < \log n \mid \mathcal{G}_t^1 \cap \mathcal{G}_t^2] = o(1) + \mathbb{P}[X_t < \log n \mid \mathcal{G}_t^1 \cap \mathcal{G}_t^2].$$

Note that $\mathbb{E}[X_t \mid \mathcal{G}_t^1 \cap \mathcal{G}_t^2] \gg \log n$, by applying the Paley-Zygmund inequality and then make use of Proposition 4.9, we get the second term above is also $o(1)$. The estimates are uniform for all $1 \leq t \leq s-1$ and the result follows. \square

To show the desired result, we cannot condition on the full information of \mathcal{F}_{s-1} . Instead, we turn to $\mathcal{I}_{s-1} = \sigma\{\mathbf{M}_{s-1}, \pi(\mathbf{M}_{s-1}), \text{MT}_t, 1 \leq t \leq s-1\}$. To the contrary of \mathcal{F}_{s-1} , the information of $\bigcup_{1 \leq t \leq s-1} \text{EXP}_t$ is not fully contained in \mathcal{I}_{s-1} . Recall the set \mathbf{M}_s in Algorithm 1 and note that \mathbf{M}_s can be determined from \mathcal{I}_{s-1} . In addition, by our choice $\kappa_0 > 4\zeta/\eta$ and the fact that each MT_t contains no more than 2ζ elements in $[n]$, it holds deterministically that $|\mathbf{M}_s| \geq \eta n/2$. Denote $V_R = \{v_i : i \in \mathbf{R}_{s-1}\}$ (recall $\mathbf{R}_{s-1} = [n] \setminus \mathbf{M}_{s-1}$) and $V_{\mathbf{M}} = \{v_i : i \in \mathbf{M}_s\}$. For any $v \in V$, let $N_R(v)$ and $N_{\mathbf{M}}(v)$ be the numbers of neighbors of v in V_R and $V_{\mathbf{M}}$, respectively. The next lemma describes the properties of the conditional distribution of G given \mathcal{I}_{s-1} , which will be useful later.

Lemma 4.21. *Recall E_0 as in (4.1). For any realization of \mathcal{I}_{s-1} , the graph G conditioned on \mathcal{I}_{s-1} is given by $GO_{s-1} \cup GO_{s-1}^\dagger$, where GO_{s-1}^\dagger is a graph on V with edges in $E_0 \setminus E(GO_{s-1})$ which is stochastically dominated by an Erdős-Rényi graph on $(V, E_0 \setminus E(GO_{s-1}))$ with edge density p_η (i.e., each edge in $E_0 \setminus E(GO_{s-1})$ is preserved with probability p_η).*

Proof. It is clear that $GO_{s-1} \subset G$ under such conditioning, and thus it remains to understand the behavior for the remaining graph GO_{s-1}^\dagger . For any $e \in E_0 \setminus E(GO_{s-1})$ and any realization $\omega_{\setminus e}$ for edges in $E_0 \setminus E(GO_{s-1})$ except e , note that if $\omega_{\setminus e}$ together with $e \in G$ (as well as $GO_{s-1} \subset E$) yields the realization \mathcal{I}_{s-1} , then $\omega_{\setminus e}$ together with $e \notin E$ also yields the realization \mathcal{I}_{s-1} . Therefore, $\mathbb{P}[e \notin E \mid \mathcal{I}_{s-1}, \omega_{\setminus e}] \geq 1 - p_\eta$, completing the proof. \square

In what follows, we fix a realization of \mathcal{I}_{s-1} .

Definition 4.22. For a triple $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ with $\mathbf{R} \subset V(\mathbf{H})$, $\mathbf{v} \in V(\mathbf{H}) \setminus \mathbf{R}$ and a vertex $v \in V_R = \{v_i : i \in [n] \setminus M_{s-1}\}$, we say a subgraph $H \subset G$ is an $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ -attaching graph rooted at v if there is an isomorphism $\phi : \mathbf{H} \rightarrow H$, such that $\phi(\mathbf{R}) \subset M_{s-1}$, $\phi(\mathbf{v}) = v$, and any two vertices in $\phi(\mathbf{R})$ has graph distance at most ζ on GO_{s-1} .

Lemma 4.23. There exists a large constant $\kappa = \kappa(\eta, \alpha_\eta, \mathbf{T}, \kappa_0)$ such that $\mathbb{P}[\mathcal{G}_\kappa \mid \mathcal{I}_{s-1}] = 1 - o(1)$ where $\mathcal{G}_\kappa = \mathcal{G}_{\kappa, s}$ is the following event: for any triple $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ where (\mathbf{R}, \mathbf{H}) is a dense pattern with \mathbf{H} being a subtree of \mathbf{T} and $\mathbf{R} = V(\mathbf{H}) \cap \mathbf{L}$, and for any vertex $v \in V_R$, the number of $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ -attaching graphs rooted at v is bounded by κ .

Proof. The proof is similar to that of Lemma 4.19, and we begin with introducing $P_\zeta(K)$ as an analogue of $P(K)$ in the proof of Lemma 4.19. For a subgraph $K \subset G$, we draw an adjunctive edge between any pair of vertices $u, v \in V(K)$ if and only if u, v has graph distance at most ζ on GO_{s-1} . This gives an adjunctive graph on $V(K)$, and we denote the collection of its components by $\mathfrak{C}_\zeta(K)$. Then we define $P_\zeta(K) = |\mathfrak{C}_\zeta(K)| - \alpha_\eta |E(K) \setminus E(GO_{s-1})|$.

Our proof is essentially by contradiction, that is, we will show that if \mathcal{G}_κ fails then a rare event must occur. To this end, let $H_1, \dots, H_\kappa \cong \mathbf{H}$ be distinct $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ -attaching graphs rooted at some $v \in V_R$, let $\phi_i : \mathbf{H} \rightarrow H_i$ be the isomorphism as in Definition 4.22 and let $K_i = H_1 \cup \dots \cup H_i$ for $1 \leq i \leq \kappa$. We claim that for $1 \leq i < \kappa$,

$$P_\zeta(K_{i+1}) \leq P_\zeta(K_i) - \delta \mathbf{1}_{\{K_{i+1} \not\subset K_i \cup GO_{s-1}\}}, \quad (4.70)$$

where $\delta > 0$ is a constant which does not depend on i or κ . We now fix i and prove (4.70). To this end, we consider components C_1, \dots, C_r of $K_{i+1} \setminus_\star K_i$ (recall the definition of $H_1 \setminus_\star H_2$ for two simple graphs as in the proof of Lemma 4.6, and we view $C_j, 1 \leq j \leq r$ as subgraphs of H_{i+1}). Let N_j be the number of components in $\mathfrak{C}_\zeta(K_{i+1})$ intersecting C_j but *not* containing v , and let

$E_j = |E(C_j) \setminus E(GO_{s-1})|$. Then, we have

$$P_\zeta(K_{i+1}) - P_\zeta(K_i) \leq \sum_{j=1}^r (N_j - \alpha_\eta E_j).$$

For each C_j , let F_j be the subgraph on vertices $V(H_{i+1}) \setminus (V(C_j) \setminus (V(K_i) \cup V(GO_{s-1})))$ with edges $E(H_{i+1}) \setminus (E(C_j) \setminus (E(K_i) \cup E(GO_{s-1})))$. We further write $F_j = F_{j,0} \cup T_{j,1} \cup \dots \cup T_{j,k_j}$, where $F_{j,0}$ is the union of components of F_j which intersect $\phi_{i+1}(\mathbf{R})$, and $T_{j,l}$'s (for $l = 1, \dots, k_j$) are the remaining tree components (here possibly $k_j = 0$). Writing $\mathbf{F}_{j,0} = \phi_{i+1}^{-1}(F_{j,0})$, we have

$$E_j = |E(H_{i+1})| - |E(F_{j,0})| - \sum_{l=1}^{k_j} |E(T_{j,l})| = |E(\mathbf{H}) \setminus E(\mathbf{F}_{j,0})| - \sum_{l=1}^{k_j} |E(T_{j,l})|. \quad (4.71)$$

For the estimation of N_j , we claim that

$$N_j \leq k_j + |V(H_{i+1})| - |V(F_{j,0})| - \sum_{l=1}^{k_j} |V(T_{j,l})| = |V(\mathbf{H}) \setminus V(\mathbf{F}_{j,0})| - \sum_{l=1}^{k_j} |E(T_{j,l})|. \quad (4.72)$$

Indeed, since $V(T_{j,l})$ belongs to a single component in $\mathfrak{C}_\zeta(K_{i+1})$ for $1 \leq l \leq k_j$ and also the vertices in $F_{j,0}$ are in a single component in $\mathfrak{C}_\zeta(K_{i+1})$ (by the definition of $(\mathbf{R}, \mathbf{H}, \mathbf{v})$ -attaching and the fact that $\phi_{i+1}(\mathbf{R}) \subset V(F_{j,0})$), we get that the number of components in $\mathfrak{C}_\zeta(K_{i+1})$ which intersect C_j is at most $1 + k_j + |V(H_{i+1})| - |V(F_{j,0})| - \sum_{l=1}^{k_j} |V(T_{j,l})|$ (this is because, the number of components on $V(F_j)$ is at most $1 + k_j$ and each vertex outside $V(F_j)$ induces at most one component). Moreover, since N_j does not count the component in $\mathfrak{C}_\zeta(K_{i+1})$ which contains v , we can remove one of the above components (possibly $F_{j,0}$ or one of the $T_{j,l}$'s) when counting N_j . This verifies (4.72). Combined with (4.71), it yields that

$$N_j - \alpha_\eta E_j \leq |V(\mathbf{H}) \setminus V(\mathbf{F}_{j,0})| - \alpha_\eta |E(\mathbf{H}) \setminus E(\mathbf{F}_{j,0})| - (1 - \alpha_\eta) \sum_{l=1}^{k_j} |E(T_{j,l})|,$$

which is $\leq -\delta \mathbf{1}_{\{\mathbf{F}_j \neq \mathbf{H}\}}$ for some $\delta > 0$ since (\mathbf{R}, \mathbf{H}) is a dense pattern (recall (4.66) and $\mathbf{R} \subset V(\mathbf{F}_{j,0})$). Letting $\mathbf{F}_j = \phi_{i+1}^{-1}(F_j)$, we then conclude (4.70) by the observation that

$$\{K_{i+1} \not\subset K_i \cup GO_{s-1}\} \subset \bigcup_{j=1}^r \{\mathbf{F}_j \neq \mathbf{H}\}, \text{ and thus } \mathbf{1}_{\{K_{i+1} \not\subset K_i \cup GO_{s-1}\}} \leq \sum_{j=1}^r \mathbf{1}_{\{\mathbf{F}_j \neq \mathbf{H}\}}.$$

In addition, for each i , we claim that there exists a large constant N depending only on the size of K_i , such that $K_{j+1} \subset K_j \cup GO_{s-1}$ cannot hold simultaneously for all $j \in \{i, i+1, \dots, i+N\}$. We prove this by contradiction. Suppose otherwise and then we see H_{i+1}, \dots, H_{i+N} must all be

contained in $K_i \cup GO_{s-1}$. Furthermore, since each H_j is connected, these graphs must be contained in the ζ -neighborhood of v in $K_i \cup GO_{s-1}$. Note that the number of vertices in this ζ -neighborhood is at most $\sum_{j=0}^{\zeta} (\Delta(GO_{s-1}))^j$ where $\Delta(GO_{s-1})$ is the maximal degree and is uniformly bounded. Recalling (4.33), we arrive at a contradiction if N is chosen sufficiently large.

Combining the preceding claim and (4.70), we can choose $\kappa = \kappa(\zeta, N, \delta, \chi)$ sufficiently large such that that $P_{\zeta}(K_{\kappa}) < -1$ on \mathcal{G}_{κ}^c . We next bound the enumeration for K_{κ} given \mathcal{I}_{s-1} . To this end, we note that for each component in $\mathfrak{C}_{\zeta}(K_{\kappa})$ the number of choices is $O(n)$ (where the O -term depends on (ζ, N, δ, χ)); this is because for each such component once a vertex is fixed the number of choices for remaining vertices is $O(1)$ by connectivity and by (4.33). Then in light of Lemma 4.21 (i), we can show that $\mathbb{P}[\mathcal{G}_{\kappa}^c] \rightarrow 0$ via a union bound over all possible choices of K_{κ} (which is similar to that for (4.67)). This completes the proof. \square

Now we are ready to present the proof of Proposition 4.10.

Proof of Proposition 4.10. Assume (4.68). Recall \mathcal{G} as in Definition 4.18 and recall $\mathcal{G}_{\kappa} = \mathcal{G}_{\kappa, s}$ from the statement of Lemma 4.23. For \mathcal{G}_s^1 , recall Definition 4.14 and the notations therein. We claim that $\mathcal{G} \cap \mathcal{G}_{\kappa} \subset \mathcal{G}_s^1$. Provided with this, we obtain from Lemmas 4.19 and 4.23 that

$$\mathbb{P}[(\mathcal{G}_s^1)^c] \leq \mathbb{P}[\mathcal{G}^c] + \mathbb{P}[(\mathcal{G}_{\kappa})^c] = o(1) + \mathbb{E}\left[\mathbb{P}[(\mathcal{G}_{\kappa})^c \mid \mathcal{I}_{s-1}]\right] = o(1).$$

We next prove the claim. To this end, for each fixed pair $(\mathbf{F}_1, \mathbf{F}_2) = (\mathbf{T}_0, \mathbf{T}_1 \cup \dots \cup \mathbf{T}_l) \in \mathcal{P}_0$, denote $\mathbf{R}_i = V(\mathbf{T}_i) \cap \mathbf{L}$ for $0 \leq i \leq l$. From the definition of \mathcal{P}_0 , we see \mathbf{T}_i is a dense tree and thus $(\mathbf{R}_i, \mathbf{T}_i)$ is a dense pattern for each $1 \leq i \leq l$.

Fix an arbitrary subgraph $\mathbf{T} \cong \mathbf{T}$ in \mathbf{G} with leaf set \mathbf{L} and $V(\mathbf{T}) \cap \mathbf{M}_{s-1} = \mathbf{L}$. Recall the definition of $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$ (with respect to \mathbf{T}) below (4.43): it counts the number of ξ -tuples $L' \in \bigcup_{1 \leq t \leq s-1} \text{EXP}_t$ which satisfy that there are two embeddings $\phi_1 : \mathbf{T}_0 \rightarrow \mathbf{T} \cup \mathbf{GO}_{s-1}$ and $\phi_2 : \mathbf{T}_1 \cup \dots \cup \mathbf{T}_l \rightarrow \mathbf{T} \cup \mathbf{GO}_{s-1}$ such that

$$\phi_1(E(\mathbf{T}_0)) \cap E(\mathbf{T}) \neq \emptyset \text{ and } \phi_1(\mathbf{R}_0) \cup \phi_2(\mathbf{R}_1 \cup \dots \cup \mathbf{R}_l) = I_G(L').$$

Note that $GO_{s-1} \cong \mathbf{GO}_{s-1}$ through $\Pi \circ I_G^{-1}$. Combined with the fact that $L' \in \bigcup_{1 \leq t \leq s-1} \text{EXP}_t$ implies $L' \bowtie_G [n]$, it yields that each such tuple corresponds to an embedding $\psi : \mathbf{T} \rightarrow G$ which satisfies the following conditions:

- (i) $\psi(\mathbf{R}_0)$ is contained in the ζ -neighborhood of $L = I_G \circ \Pi^{-1}(\mathbf{L})$ on GO_{s-1} ;

(ii) for each $1 \leq i \leq l$, we have $\psi(\mathbf{R}_i) \subset GO_{s-1}$ and the diameter of $\psi(\mathbf{R}_i)$ with respect to the graph metric on GO_{s-1} is at most ζ .

Thus, we can bound $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$ by the number of such embeddings. To this end, note that on \mathcal{G} we have

$$\Delta(G) \leq \kappa_4 n p_\eta. \quad (4.73)$$

For the number of possible realizations of $\psi(V(\mathbf{T}_0))$, when $\mathbf{R}_0 = \emptyset$ it is bounded by $O(n \times (np_\eta)^{|E(\mathbf{T}_0)|})$ using (4.73); when $\mathbf{R}_0 \neq \emptyset$ it is bounded by $O(\mathbb{D}(\mathbf{T}_0))$ using (4.33) and (4.63) (More precisely, the number of ways of choosing $\psi(\mathbf{R}_0)$ is $O(1)$ by (4.33) and for each fixed $\psi(\mathbf{R}_0)$, the number of ways of choosing the rest of $\psi(V(\mathbf{T}_0))$ is $O(\mathbb{D}(\mathbf{T}_0))$ by (4.63)). In addition, for any edge $(\mathbf{u}, \mathbf{v}) \in E(\mathbf{T}) \setminus (E(\mathbf{T}_0) \cup \dots \cup E(\mathbf{T}_l))$, once $\psi(\mathbf{u})$ is fixed, the number of choices for $\psi(\mathbf{v})$ is $O(np_\eta)$ by (4.73). For each $1 \leq i \leq l$ and any $\mathbf{v} \in V(\mathbf{T}_i)$, once $\psi(\mathbf{v})$ is fixed, each realization of $\psi(V(\mathbf{T}_i))$ corresponds to a $(\mathbf{R}_i, \mathbf{T}_i, \mathbf{v})$ -attaching graph rooted at $\psi(\mathbf{v})$, and thus the number of such realizations is at most κ by \mathcal{G}_κ .

Provided with preceding observations, we may bound the number of these embeddings (which in turn bounds $\text{Enum}(\mathbf{F}_1, \mathbf{F}_2)$) by first choosing $\psi(V(\mathbf{T}_0))$ and then choosing the ψ -values for the remaining vertices on \mathbf{T} inductively. More precisely, the ordering for choosing vertices satisfy the following properties (see Figure 4.3 for an illustration): (i) we first determine the ψ -value of vertices in $V(\mathbf{T}_0)$; (ii) for the remaining vertices, the vertex whose ψ -value is to be chosen is neighboring to some vertex whose ψ -value has been chosen; (iii) whenever we choose $\psi(\mathbf{v})$ for any $\mathbf{v} \in V(\mathbf{T}_i)$, we also choose $\psi(V(\mathbf{T}_i))$ immediately (which has at most κ choices as noted above). Therefore, we conclude that

$$\text{Enum}(\mathbf{F}_1, \mathbf{F}_2) \lesssim \begin{cases} n \times (np_\eta)^{|E(\mathbf{T}_0)|} \times (np_\eta)^{|E(\mathbf{T}) \setminus (E(\mathbf{T}_0) \cup \dots \cup E(\mathbf{T}_l))|}, & \text{if } \mathbf{R}_0 = \emptyset, \\ \mathbb{D}(\mathbf{T}_0) \times (np_\eta)^{|E(\mathbf{T}) \setminus (E(\mathbf{T}_0) \cup \dots \cup E(\mathbf{T}_l))|}, & \text{if } \mathbf{R}_0 \neq \emptyset, \end{cases}$$

which implies (4.46) with an appropriate choice of κ_1 . This proves the claim.

Next we investigate \mathcal{G}_s^2 as in Definition 4.15. For Item (i), we first note that by the sub-Gaussian concentration of Bernoulli variables, it holds that each vertex in G has degree at least $(1 - \eta/8)np_\eta$ except with exponentially small probability. Furthermore, from Lemma 4.21 we also obtain that under any conditioning of \mathcal{I}_{s-1} , except with exponentially small probability we have for any $v \in V_R$, the total number of edges between v and vertices in $V \setminus V_M$ is no more than $(1 + \eta/8)(n - |M_s|)p_\eta$.

two tuples $L = I_G^{-1}(L) = (l_1, \dots, l_\xi)$ and $L^* = (l_1^*, \dots, l_\xi^*)$ must satisfy the following two properties:

- (i) For any $i \in [\xi]$, there exists $j \in [\xi]$ such that $d_{GO_{s-1}}(v_{l_i^*}, v_{l_j}) \leq \zeta$.
- (ii) For at least two indices $p \in [\xi]$, there exists $q \in [\xi]$ such that $d_{GO_{s-1}}(v_{l_p}, v_{l_q^*}) \leq \zeta$.

We denote by $L(L)$ the collection of all ξ -tuples $L^* \in \mathfrak{A}(M_{s-1}, \xi)$ that satisfy Property (i) above.

For any triple (L, Q, L^*) with $L^* \in L(L)$, we divide the event $L^* \in \text{EXP}_{t^*}$ for some $1 \leq t^* \leq s-1$ into two cases according to whether

$$t^* \in \mathfrak{t}(L) \text{ where } \mathfrak{t}(L) = \{t' : d_{GO_{s-1}}(v_{u_{t'}}, v_{l_i}) \leq \zeta \text{ for some } i \in [\xi]\}. \quad (4.76)$$

For the case that $t^* \notin \mathfrak{t}(L)$, the tree that certifies $L' \in \text{CAND}_{t^*}$ can not be fully contained in $T_{L,Q} \cup GO_{s-1}$. From Lemma 4.21 and the proof of Lemma 4.11 (see e.g., (4.39)), we see such case happens with probability $o(1)$ uniformly under any conditioning $\mathcal{I}_{s-1} \cup \{T_{L,Q} \subset G\}$. For the case that $t^* \in \mathfrak{t}(L)$, we apply a union bound and thus we get

$$\begin{aligned} \mathbb{P}[L \bowtie_G Q, L \notin \text{GC}_s \mid \mathcal{I}_{s-1}] &\leq o(1) \times \mathbb{P}[L \bowtie_G Q \mid \mathcal{I}_{s-1}] \\ &+ \sum_{L^* \in L(L)} \sum_{t^* \in \mathfrak{t}(L)} \mathbb{P}[L \bowtie_G Q, L^* \in \text{EXP}_{t^*} \mid \mathcal{I}_{s-1}]. \end{aligned}$$

From Lemma 4.21 we see $\mathbb{P}[L \bowtie_G Q \mid \mathcal{I}_{s-1}] \leq p_\eta^\zeta$ for any realization of \mathcal{I}_{s-1} . Thus the first term above is $o(p_\eta^\zeta)$. To treat the remaining terms, we note that

- (a) $L^* \in \text{EXP}_{t^*}$ implies that $L^* \in \bigcup_{1 \leq t \leq s-1} \text{SUC}_t$ or $L^* \in \text{FAIL}_{t^*} \setminus \bigcup_{1 \leq t \leq s-1} \text{SUC}_t$;
- (b) if $L^* \in \bigcup_{1 \leq t \leq s-1} \text{SUC}_t$, then from Property (ii) above,

$$\text{there are two indices } i, j \in L \text{ such that } d_{GO_{s-1}}(v_i, v_j) \leq 3\zeta; \quad (4.77)$$

- (c) $\{L^* \in \text{FAIL}_{t^*}\} = \{I_{t^*} = 0\} \cup \{I_{t^*} = 1, L^* \prec L_{t^*}\}$ (recall that $\text{MT}_t = (L_t, Q_t, Q'_t)$ for $1 \leq t \leq s-1$ with $I_t = 1$).

Combining things together, we see for each pair (L, Q) , any $L^* \in L(L)$ and any $t^* \in \mathfrak{t}(L)$, it holds that $\mathbb{P}[L \bowtie_G Q, L^* \in \text{EXP}_{t^*} \mid \mathcal{I}_{s-1}]$ is bounded by p_η^ζ times

$$\begin{aligned} &\mathbf{1}_{\{L \text{ satisfies (4.77)}\}} + \mathbb{P}[X_t < \log n \mid \mathcal{I}_{s-1}, L \bowtie_G Q] \\ &+ \mathbb{P}[X_{t^*} \geq \log n, L^* \notin \bigcup_{1 \leq t \leq s-1} \text{SUC}_t, L^* \prec L_{t^*} \mid \mathcal{I}_{s-1}, L \bowtie_G Q]. \end{aligned} \quad (4.78)$$

In order to bound the second term in (4.78), note that if $\{X_t < \log n\}$ holds under $\mathcal{I}_{s-1} \cap \{L \bowtie_G Q\}$, then it also holds under \mathcal{I}_{s-1} together with any other configuration on the set $E(T_{L,Q})$. Thus,

$$\mathbb{P}[X_{t^*} < \log n \mid \mathcal{I}_{s-1}, L \bowtie_G Q] \leq \mathbb{P}[X_{t^*} < \log n \mid \mathcal{I}_{s-1}].$$

In order to bound the third term in (4.78), denote Avai_t for the (random) set of tuples in CAND_t that can be successfully matched, $1 \leq t \leq s-1$. Then $|\text{Avai}_t| = X_t$ and the conditional law of \prec given $\mathcal{I}_{s-1} \cap \{L \bowtie_G Q\}$ is uniform conditioned on the following event:

$$\{L_t \text{ is } \prec\text{-minimal among } \text{Avai}_t, \text{ for all } 1 \leq t \leq s-1 \text{ with } I_t = 1\}.$$

In particular, under such conditioning, for $L^* \notin \bigcup_{1 \leq t \leq s-1} \text{SUC}_t$ it holds with conditional probability $1/X_{t^*}$ that $L^* \prec L_{t^*}$. As a result, the third term is bounded by $(\log n)^{-1} = o(1)$.

Summing over all (L, Q) and combining all the arguments above altogether, we get

$$\begin{aligned} & \sum_{L \in \mathfrak{A}(\mathbb{M}_s, \xi)} \mathbb{P}[L \in \text{CAND}_s \setminus \text{GC}_s \mid \mathcal{I}_{s-1}] \\ & \leq \sum_{L \in \mathfrak{A}(\mathbb{M}_s, \xi)} \sum_{Q \in \mathfrak{A}_s} \sum_{L^* \in \mathcal{L}(L)} \sum_{t^* \in \mathfrak{t}(L)} p_\eta^\zeta \left(o(1) + \mathbf{1}_{\{L \text{ satisfies (4.77)}\}} + \mathbb{P}[X_{t^*} < \log n \mid \mathcal{I}_{s-1}] \right). \end{aligned} \quad (4.79)$$

We now bound the number of effective terms in the summation of (4.79). Clearly, $|\mathfrak{A}(\mathbb{M}_t, \xi)| \leq n^\xi$ and $|\mathfrak{A}_s| \leq n^{\chi-1}$. In addition, by (4.33) we see both $|\mathcal{L}(L)|$ and $|\mathfrak{t}(L)|$ are uniformly bounded for $L \in \mathfrak{A}(\mathbb{M}_s, \xi)$. Furthermore, the number of $L \in \mathfrak{A}(\mathbb{M}_s, \xi)$ that satisfy (4.77) is of order $O(n^{\xi-1})$, and for each $1 \leq t^* \leq s-1$ the number of tuples $L \in \mathfrak{A}(\mathbb{M}_s, \xi)$ such that $\mathfrak{t}(L)$ contains t^* is also of order $O(n^{\xi-1})$ (since some vertex in L must be close to v_{u_t} on the graph GO_{s-1}). Altogether, we see the right hand side of (4.79) is upper-bounded by (recall that $\zeta = \xi + \chi - 1$)

$$O(n^\zeta) \times o(p_\eta^\zeta) + O(n^{\zeta-1}) \times p_\eta^\zeta + O(n^{\zeta-1}) \times p_\eta^\zeta \times \sum_{1 \leq t^* \leq s-1} \mathbb{P}[X_{t^*} < \log n \mid \mathcal{I}_{s-1}].$$

Averaging over \mathcal{I}_{s-1} and recalling (4.74) and (4.79), we obtain that

$$\mathbb{E}|\text{CAND}_s \setminus \text{GC}_s| = o((np_\eta)^\zeta) + O(n^{\zeta-1} p_\eta^\zeta) \sum_{1 \leq t^* \leq s-1} \mathbb{P}[X_{t^*} < \log n] \stackrel{(4.69)}{=} o((np_\eta)^\zeta),$$

as desired. This implies that Item (i) of \mathcal{G}_s^2 in Definition 4.15 holds with probability $1 - o(1)$.

Finally, we treat Item (ii) of \mathcal{G}_s^2 in Definition 4.15 and it suffices to show that it holds on the event \mathcal{G} . Since each tuple in GC_s corresponds to a subgraph $T \cong \mathbf{T}$ in G rooted at v_{u_s} , (4.73) implies that $|\text{GC}_s| = O((np_\eta)^\zeta)$. For each fixed $L_i \in \text{GC}_s$ and a nonempty subset $\mathbf{R} \subset \mathbf{L}$, we bound the number of $L_j \in \text{GC}_s$ such that $(L_i, L_j) \in \text{IP}_s(\mathbf{R})$ as follows: each such L_j corresponds to a subgraph $T \cong \mathbf{T}$ of G with leaves in \mathbf{R} mapped to a fixed subset in $V(G)$ under the isomorphism. In order to bound the enumeration for such T , we use the following two-step procedure to choose T :

- choose a subgraph $T_0 \cong \mathbf{Span}(\mathbf{R})$ of G with leaves in \mathbf{R} mapped to a fixed subset in $V(G)$ under the isomorphism;
- choose $T \cong \mathbf{T}$ of G with $\mathbf{Span}(\mathbf{R})$ mapped to T_0 under the isomorphism.

Thus, by Definition 4.18 and by (4.73), the number of choices for such T is $O(\mathbb{D}(\mathbf{Span}(\mathbf{R})) \times (np_\eta)^{|E(\mathbf{T}) \setminus E(\mathbf{Span}(\mathbf{R}))|})$ on \mathcal{G} , and thus Item (ii) of \mathcal{G}_s^2 holds with probability $1 - o(1)$. This completes the proof of Proposition 4.10. \square

References

- [1] E. Abbe. Community detection and stochastic block models: recent developments. *J. Mach. Learn. Res.*, 18:Paper No. 177, 86, 2017.
- [2] A. S. Bandeira, A. Perry, and A. S. Wein. Notes on computational-to-statistical gaps: predictions using statistical physics. *Port. Math.*, 75(2):159–186, 2018.
- [3] B. Barak, C.-N. Chou, Z. Lei, T. Schramm, and Y. Sheng. (nearly) efficient algorithms for the graph matching problem on correlated random graphs. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [4] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, volume 1, pages 26–33 vol. 1, 2005.
- [5] M. Bozorg, S. Salehkaleybar, and M. Hashemi. Seedless graph matching via tail of degree distribution for correlated erdos-renyi graphs. Preprint, arXiv:1907.06334.
- [6] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006.
- [7] D. Cullina and N. Kiyavash. Exact alignment recovery for correlated Erdos-Rényi graphs. Preprint, arXiv:1711.06783.
- [8] D. Cullina and N. Kiyavash. Improved achievability and converse bounds for erdos-renyi graph matching. In *Proceedings of the 2016 ACM SIGMETRICS International Conference on*

- Measurement and Modeling of Computer Science*, SIGMETRICS '16, page 6372, New York, NY, USA, 2016. Association for Computing Machinery.
- [9] D. Cullina, N. Kiyavash, P. Mittal, and H. V. Poor. Partial recovery of erdos-rényi graph alignment via k -core alignment. SIGMETRICS '20, page 99100, New York, NY, USA, 2020. Association for Computing Machinery.
 - [10] O. E. Dai, D. Cullina, N. Kiyavash, and M. Grossglauser. Analysis of a canonical labeling algorithm for the alignment of correlated erdos-rényi graphs. *Proc. ACM Meas. Anal. Comput. Syst.*, 3(2), jun 2019.
 - [11] A. Dembo, A. Montanari, and S. Sen. Extremal cuts of sparse random graphs. *Ann. Probab.*, 45(2):1190–1217, 2017.
 - [12] J. Ding and H. Du. Detection threshold for correlated erdos-renyi graphs via densest subgraph. Preprint, arXiv:2203.14573.
 - [13] J. Ding and H. Du. Matching recovery threshold for correlated random graphs. Preprint, arXiv:2205.14650.
 - [14] J. Ding, Z. Ma, Y. Wu, and J. Xu. Efficient random graph matching via degree profiles. *Probab. Theory Related Fields*, 179(1-2):29–115, 2021.
 - [15] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations II: Erdos-rényi graphs and universality. Preprint, arXiv:1907.08883.
 - [16] Z. Fan, C. Mao, Y. Wu, and J. Xu. Spectral graph matching and regularized quadratic relaxations: Algorithm and theory. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2985–2995. PMLR, 13–18 Jul 2020.
 - [17] S. Feizi, G. Quon, M. Medard, M. Kellis, and A. Jadbabaie. Spectral alignment of networks. Preprint, arXiv:1602.04181.
 - [18] D. Gamarnik. The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences*, 118(41):e2108492118, 2021.
 - [19] D. Gamarnik and I. Zadik. The landscape of the planted clique problem: Dense subgraphs and the overlap gap property. Preprint, arXiv:1904.07174.

- [20] L. Ganassali and L. Massoulié. From tree matching to sparse graph alignment. In J. Abernethy and S. Agarwal, editors, *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1633–1665. PMLR, 09–12 Jul 2020.
- [21] A. Haghighi, A. Ng, and C. Manning. Robust textual inference via graph matching. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 387–394, Vancouver, British Columbia, Canada, Oct 2005.
- [22] G. Hall and L. Massoulié. Partial recovery in the graph alignment problem. Preprint, arXiv:2007.00533.
- [23] E. Kazemi, S. H. Hassani, and M. Grossglauser. Growing a graph matching from a handful of seeds. *Proc. VLDB Endow.*, 8(10):10101021, jun 2015.
- [24] D. Kunisky, A. S. Wein, and A. S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. Preprint, arXiv:1907.11636.
- [25] V. Lyzinski, D. E. Fishkind, and C. E. Priebe. Seeded graph matching for correlated Erdos-Rényi graphs. *J. Mach. Learn. Res.*, 15:3513–3540, 2014.
- [26] P. Manurangsi, A. Rubinfeld, and T. Schramm. The Strongish Planted Clique Hypothesis and Its Consequences. In J. R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:21, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.
- [27] C. Mao, M. Rudelson, and K. Tikhomirov. Exact matching of random graphs with constant correlation. Preprint, arXiv:2110.05000.
- [28] C. Mao, Y. Wu, J. Xu, and S. H. Yu. Testing network correlation efficiently via counting trees. Preprint, arXiv:2110.11816.
- [29] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, USA, 2005.
- [30] A. Montanari. Optimization of the sherrington-kirkpatrick hamiltonian. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1417–1433, 2019.

- [31] A. Montanari and S. Sen. Semidefinite programs on sparse random graphs and their application to community detection. In *STOC'16—Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 814–827. ACM, New York, 2016.
- [32] E. Mossel and J. Xu. Seeded graph matching via large neighborhood statistics. *Random Structures Algorithms*, 57(3):570–611, 2020.
- [33] A. Narayanan and V. Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, 2008.
- [34] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187, 2009.
- [35] R. Nenadov, A. Steger, and P. Su. An $O(N)$ time algorithm for finding Hamilton cycles with high probability. In *12th Innovations in Theoretical Computer Science Conference*, volume 185 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 60, 17. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2021.
- [36] P. Pedarsani and M. Grossglauser. On the privacy of anonymized networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, page 12351243, New York, NY, USA, 2011. Association for Computing Machinery.
- [37] P. Raghavendra, T. Schramm, and D. Steurer. High dimensional estimation via sum-of-squares proofs. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 3389–3423. World Sci. Publ., Hackensack, NJ, 2018.
- [38] F. Shirani, S. Garg, and E. Erkip. Seeded graph matching: Efficient algorithms and theoretical guarantees. In *2017 51st Asilomar Conference on Signals, Systems, and Computers*, pages 253–257, 2017.
- [39] R. Singh, J. Xu, and B. Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences of the United States of America*, 105:12763–8, 10 2008.
- [40] J. Spencer. Counting extensions. *J. Combin. Theory Ser. A*, 55(2):247–255, 1990.
- [41] E. Subag. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.*, 74(5):1021–1044, 2021.

- [42] J. T. Vogelstein, J. M. Conroy, V. Lyzinski, L. J. Podrazik, S. G. Kratzer, E. T. Harley, D. E. Fishkind, R. J. Vogelstein, and C. E. Priebe. Fast approximate quadratic programming for graph matching. *PLOS ONE*, 10(4):1–17, 04 2015.
- [43] Y. Wu and J. Xu. *Statistical Problems with Planted Structures: Information-Theoretical and Computational Limits*, page 383424. Cambridge University Press, 2021.
- [44] Y. Wu, J. Xu, and S. H. Yu. Settling the sharp reconstruction thresholds of random graph matching. Preprint, arXiv:2102.00082.
- [45] Y. Wu, J. Xu, and S. H. Yu. Testing correlation of unlabeled random graphs. Preprint, arXiv:2008.10097.
- [46] L. Yartseva and M. Grossglauser. On the performance of percolation graph matching. In *Proceedings of the First ACM Conference on Online Social Networks, COSN '13*, page 119130, New York, NY, USA, 2013. Association for Computing Machinery.
- [47] L. Zdeborová and F. Krzakala. Statistical physics of inference: thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.

5 Algorithmic phase transition of the random graph alignment problem

In this section we establish an algorithmic phase transition for the graph alignment problem involving two independent Erdős-Rényi graphs as their edge density, denoted by p , evolves. We show that in the sparse regime, polynomial-time approximation schemes exist, while in the dense regime, a $\sqrt{8/9}$ multiplicative factor gap between computationally achievable solutions and theoretical optimums emerges. By determining the precise informational and computational thresholds for both regimes, we successfully identify the critical window within which the phase transition takes place. This section is based on a joint work with Shuyang Gong and Rundong Huang.

5.1 Introduction and main results

The *Graph Alignment Problem* (GAP) for two simple graphs with the same number of vertices involves finding a vertex bijection which maximizes the size of overlap of these two graphs through such correspondence. Specifically, given $G(V, E)$ and $G(V, E)$ with $|V| = |V|$, the goal is to find an one-to-one mapping $\pi : V \rightarrow V$ that maximizes the expression

$$\sum_{v_i \neq v_j} \mathbf{1}_{(v_i, v_j) \in E} \mathbf{1}_{\pi(v_i), \pi(v_j) \in E}.$$

Closely related to the *Maximum Common Subgraph Problem* (MCS), GAP is an important but challenging combinatorial optimization problem which plays essential roles in various applied fields like computational biology [37, 38], social networking [30, 31], computer vision [3, 9] and natural language processing [20]. The study of efficient algorithms for solving GAP exactly or approximately has been done extensively through the past decades.

Unfortunately, as a special case of the *Quadratic Assignment Problem* (QAP) [32, 5], the exact solution of GAP is known to be NP-hard, which means no known algorithm that can solve it in polynomial times for all instances. Moreover, it was shown in [25] that approximating QAP within a factor $2^{\log^{1-\varepsilon}(n)}$ for any $\varepsilon > 0$ is also NP-hard, so find near-optimal solutions for GAP efficiently seems to be out of reach in general.

Given the worst-case hardness result, analysis for GAP in the existing literature are often done to a specific class of instances, such as the sparse graphs (e.g. [22, 23]) or correlated Erdős-Rényi

graphs (e.g. [4, 13, 28, 29], which we will discuss in details in Subsection 5.1.1). In this paper we consider GAP over typical instances for a pair of independent Erdős-Rényi graphs which we call it the random GAP. Our primary objective is to find near-optimal solutions within a multiplicative constant factor.

Formally we fix $n \in \mathbb{N}, p \in (0, 1)$, denote $V = \{v_1, \dots, v_n\}, \mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ and let G, \mathbf{G} be independent samples of Erdős-Rényi graphs on V, \mathbf{V} with edge density p , respectively. For $1 \leq i \neq j \leq n$, define

$$G_{ij} = G_{ji} = \mathbf{1}_{(v_i, v_j) \in E}, \quad \mathbf{G}_{ij} = \mathbf{G}_{ji} = \mathbf{1}_{(\mathbf{v}_i, \mathbf{v}_j) \in \mathbf{E}},$$

then $\{G_{ij}\}_{1 \leq i < j \leq n}, \{\mathbf{G}_{ij}\}_{1 \leq i < j \leq n}$ are independent Bernoulli variables with parameter p by definition. For any permutation $\pi \in S_n$, we define

$$O(\pi) = \sum_{1 \leq i < j \leq n} G_{ij} G_{\pi(i)\pi(j)}$$

for the overlap of G, \mathbf{G} along with the vertex correspondence $v_i \mapsto \mathbf{v}_{\pi(i)}$. Since for any $\pi \in S_n$, we have

$$\mathbb{E} O(\pi) = \binom{n}{2} p^2 \stackrel{\text{def}}{=} E_{n,p}, \quad (5.1)$$

we also consider the centered version of the family of variables

$$\{\tilde{O}(\pi)\}_{\pi \in S_n} \stackrel{\text{def}}{=} \{O(\pi) - E_{n,p}\}_{\pi \in S_n}. \quad (5.2)$$

For positive functions $f, g : \mathbb{N} \rightarrow \mathbb{R}^+$, we use standard notations like $f = o(g), O(g), \Omega(g)$ to mean f/g is converging to 0, upper bounded, lower bounded from 0, respectively. And we use $f \ll g$ (resp. $f \gg g$) to mean f/g tends to 0 (resp. ∞) as $n \rightarrow \infty$. Our first result establishes the asymptotic of the maximum centered overlap for various p .

Theorem 5.1 (Informational-threshold). *Denote $p_c = \sqrt{\log n/n}$, then the following hold.*

- (Sparse regime) *If $\log n/n \ll p \ll p_c$, then for $S_{n,p} = \frac{n \log n}{\log(\frac{\log n}{np^2})}$, we have*

$$\max_{\pi \in S_n} \tilde{O}_\pi / S_{n,p} \xrightarrow{\text{in probability}} 1, \text{ as } n \rightarrow \infty. \quad (5.3)$$

- (Dense regime) *If $p_c \ll p \ll 1/\log n$, then for $D_{n,p} = \sqrt{n^3 p^2 \log n}$, we have*

$$\max_{\pi \in S_n} \tilde{O}_\pi / D_{n,p} \xrightarrow{\text{in probability}} 1, \text{ as } n \rightarrow \infty. \quad (5.4)$$

Remark 5.2. Based on Theorem 5.1, we will adopt the following conventions going forward: $\log n/n \ll p \ll p_c$ represents the *sparse regime*, and $p_c \ll p \ll 1/\log n$ represents the *dense regime*. A simple computation reveals the following relationships:

$$p \ll p_c \Rightarrow E_{n,p} \ll S_{n,p} \quad \text{and} \quad p \gg p_c \Rightarrow E_{n,p} \gg D_{n,p}.$$

From these relationships, it follows that the typical first-order asymptotic behavior of $\max_{\pi \in S_n} O_\pi$ is governed by $S_{n,p}$ in the sparse regime, while in the dense regime, it is captured by the expectation $E_{n,p}$. Moreover, applying concentration inequalities suggest that in the dense regime, all the O_π values concentrate around $E_{n,p}$. This explains why we focus on the centered version $\tilde{O}(\pi)$. Otherwise, the problem would become trivial in the dense regime, both in terms of informational and computational aspects.

While there are further indications of the informational-threshold for the parameter p in other regimes, this paper focuses on the algorithmic phase transition between the sparse and dense regimes. Therefore, we defer the discussion about informational results in the remaining regimes and shift our attention to the computational aspect of the problem.

Denote G_n as the collection of all simple graphs on n vertices, and for any $G, G \in G_n$ sampled independently from the law of $\mathbf{G}(n, p)$, let $S_\beta(G, G)$ be the set of permutation π which satisfies $\tilde{O}(\pi) \geq \beta S_{n,p}$ (resp. $\tilde{O}(\pi) \geq \beta D_{n,p}$) for $p \ll p_c$ (resp. $p \gg p_c$). Namely, $S_\beta(G, G)$ is the set of asymptotically β -optimal solutions to the random GAP. To consider the power of algorithms for this problem, formally we make the following definition.

Definition 5.1 (Graph alignment algorithms). *For a (potentially randomized) graph alignment algorithm, we mean an algorithm $\mathcal{A} = \mathcal{A}(\Omega, f, \{\mathbb{P}_{G,G}\})$, where Ω is some abstract sample space, f is a deterministic function from $G_n \times G_n \times \Omega$ to S_n , $\{\mathbb{P}_{G,G}\}$ is a family of probability measures on Ω indexed by $G_n \times G_n$, and \mathcal{A} is defined as follows: for an input (G, G) , the output $\mathcal{A}(G, G) = f(G, G, \omega) \in S_n$ with some ω sampled from $\mathbb{P}_{G,G}$.*

Denote the set of graph alignment algorithm as GAA.

We aim to determine the power limit of *efficient algorithms* in GAA. Specifically, we seek to identify the values of β for which there exists a polynomial-time algorithm $\mathcal{A} \in \text{GGA}$ such that the probability $\mathbb{P}[\mathcal{A}(G, G) \in S_\beta(G, G)]$ is non-vanishing or close to one. Here, the probability is

taken over the random input $(G, \mathbf{G}) \sim \mathbf{G}(n, p)^{\otimes 2}$, together with the additional randomness in the algorithm itself.

Our first result states that for p in the sparse regime, polynomial-time algorithm may successfully find solutions in $S_\beta(G, \mathbf{G})$ with high probability for β arbitrarily close to 1.

Theorem 5.3 (Polynomial-time approximation scheme in the sparse regime). *For any $\varepsilon > 0$ and $\log n/n \ll p \ll p_c$, there is a polynomial-time algorithm $\mathcal{A} \in \text{GAA}$ (depending on n, p, ε) such that as $n \rightarrow \infty$,*

$$\mathbb{P}[\mathcal{A}(G, \mathbf{G}) \in S_{1-\varepsilon}(G, \mathbf{G})] = 1 - o(1).$$

According to Theorem 5.3, no statistical-computation gap exists in the sparse regime. However, in the dense regime, algorithmic barriers emerge, and our objective is to identify the precise threshold for such barriers. While lower bounds can be established by constructing and analyzing specific algorithms, determining a non-trivial upper bound for the entire class of efficient algorithms remains a challenging task. Nonetheless, there are strong indications and widely-used approaches to showcase algorithmic limitations beyond certain thresholds in specific sub-classes. In this paper, we concentrate on a particular class of algorithms referred as *online algorithms*, which are defined below.

Definition 5.2 (Online algorithms). *We define an algorithm $\mathcal{A} \in \text{GAA}$ as an online algorithm if the output π^* of \mathcal{A} satisfies the following conditions:*

- *The values of $\pi^*(1), \pi^*(2), \dots, \pi^*(n)$ are determined sequentially.*
- *For each $1 \leq k \leq n$, once $\pi^*(1), \dots, \pi^*(k-1)$ are determined, $\pi^*(k)$ is a random variable (with internal randomness of the algorithm itself) that takes a value in $[n] \setminus \{\pi^*(1), \dots, \pi^*(k-1)\}$ with a distribution \mathbb{P}_k determined by $\{G_{ij}\}_{1 \leq i < j \leq k}$, $\{\mathbf{G}_{ij}\}_{1 \leq i < j \leq n}$, and $\pi^*(1), \dots, \pi^*(k-1)$.*

We denote the set of online algorithms as OGAA.

Intuitively, one might think of G as a graph that is constructed online, vertex by vertex, while \mathbf{G} is a pre-specified offline graph. From this perspective, the above definition implies that an online algorithm must determine $\pi^*(k)$ using only the information available after the k -th update of the vertices in G . The following result illustrates a sharp transition in the performance of online algorithms in the dense regime.

Theorem 5.4 (Statistical-computation gap in the dense regime). *Denote $\beta_c = \sqrt{8/9}$, then for any $\epsilon > 0$ and $p_c \ll p \ll 1$, the following hold.*

- (i) *There exists a deterministic algorithm $\mathcal{A}^* \in \text{OGGA}$ which runs in $O(n^3)$ time, such that as $n \rightarrow \infty$,*

$$\mathbb{P}[(G, \mathbb{G}) : \mathcal{A}^*(G, \mathbb{G}) \in S_{\beta_c - \epsilon}(G, \mathbb{G})] = 1 - o(1).$$

In other words, \mathcal{A}^ finds $(\beta_c - \epsilon)$ -optimal solutions for typical instances (G, \mathbb{G}) .*

- (ii) *There exists $c = c(\epsilon) > 0$ and a set of graphs \mathcal{H} such that $\mathbb{P}[\mathbb{G} \in \mathcal{H}] = 1 - o(1)$ as $n \rightarrow \infty$, and for any $\mathbb{G} \in \mathcal{H}$ and any $\mathcal{A} \in \text{OGAA}$,*

$$\mathbb{P}\left[G : \mathbb{Q}[\mathcal{A}(G, \mathbb{G}) \in S_{\beta_c + \epsilon}(G, \mathbb{G})] \geq \exp(-cn \log n)\right] \leq \exp(-cn \log n),$$

where \mathbb{Q} denotes the internal randomness of the algorithm \mathcal{A} . In other words, no online algorithm can find $(\beta_c + \epsilon)$ -optimal solutions with probability exceeding $\exp(-\Omega(n \log n))$ under typical instances of (G, \mathbb{G}) .

Remark 5.5. The definition of online algorithms can be further generalized to allow for the matching of any $o(n)$ pairs of vertices in each round. With this modification, the hardness result can still be obtained with only minor changes in the wording. Furthermore, the probability upper bound $\exp(-\Omega(n \log n))$ is exponentially tight up to a multiplicative constant, since even trivial algorithms like random guess will be able to find solutions in $S_{\beta_c + \epsilon}(G, \mathbb{G})$ with probability at least $\exp(-n \log n)$ under typical instances of (G, \mathbb{G}) .

In addition, it is anticipated that our framework can demonstrate the failure of any sufficiently stable algorithm in GAA to find solutions beyond the threshold β_c with a non-vanishing probability. Determining the precise threshold for stable algorithms remains an intriguing avenue for future research.

Theorem 5.4 provides a characterization for the computational threshold for the random GAP in the dense regime, demonstrating the existence of a multiplicative constant factor gap of $\beta_c = \sqrt{8/9}$ between theoretical optimal solutions and computationally available solutions. Comparing to the sparse regime, where no statistical-computation gap emerges, an algorithmic phase transition of the random GAP occurs within the critical window $p \asymp p_c$. Notably, for $p = Cp_c$ with constant $C \in (0, \infty)$, it can be argue that the computation to information ratio of the random GAP with

parameter p continuously interpolates between 1 and β_c as C evolves from 0 to ∞ . This implies the algorithmic phase transition is a coarse one, with a critical window in length comparable to the critical point p_c . Our findings align with the observation in [24] that coarse transitions usually indicate low complexity.

5.1.1 Backgrounds and prior works

Our motivation for considering the random GAP is twofold. On one hand, the GAP for a pair of independent Erdős-Rényi graphs arises naturally in the study of correlated random graph models, which have been extensively explored in the field of combinatorial statistics in recent years. A deeper understanding of either the informational threshold or the computational threshold for random GAP will also shed lights on the correlated model itself. On the other hand, there is inherent intrigue in gaining insights into the computational complexity of such a natural and important problem. Although the computational intractability for the worst-case scenario is already known (primarily due to certain highly structured "bad" instances), the situation for the average-case setting remains unclear.

We now delve into these two aspects that motivate our current work in greater detail.

Correlated random graph model The *correlated random graph model* refers to a pair of correlated Erdős-Rényi random graphs with the same number of vertices, where the correlation between them is determined by a *latent* vertex bijection. The study of the correlated model primarily focuses on recovering the hidden correspondence based solely on the topological structures of these two graphs, as well as the closely related correlation detection problem. Significant progress has been made in recent years, including information-theoretic analysis [8, 7, 19, 41, 42, 10, 11] and proposals for various efficient algorithms [4, 13, 28, 29]. Currently, the community has achieved a comprehensive understanding of the informational thresholds for the correlated random graph model. However, a substantial statistical-computational gap remains, and the precise computational threshold is still undetermined.

The initial exploration of the GAP over two independent instances of Erdős-Rényi graphs stemmed from the study of correlation detection for a pair of random graphs. This can be formulated as a hypothesis testing problem, where, under the null hypothesis, the two graphs are independently sampled, while under the alternative hypothesis, the pair is sampled from the correlated law. The authors of [41] were the first to investigate this problem, and they introduced the maximal overlap

of these two graphs as the testing statistic, which is where the concept of random GAP arises. However, the authors only derived an informational upper bound using a straightforward first moment method, which was sufficient for their purposes. Subsequently, in [15, Section 7, Open question Q1], the authors formally formulated the random GAP problem and sought to determine the exact informational thresholds for different regimes of p . One of the main objectives of this paper is to provide an answer to this question.

Computational complexity of random optimization problems *Random optimization problems* often involve optimizing an objective function generated from random data. The computational complexity of random optimization problems in high-dimensional settings is currently an active and challenging research area. As mentioned earlier, while worst-case hardness results are well-established, the average-case complexity of these problems can be intricate.

In general, a statistical-computation gap may exist, indicating that efficient algorithms encounter barriers below the optimal threshold. Alternatively, there might be a polynomial-time approximation scheme that finds near-optimal solutions for typical instances of the random input. The presence of such statistical-computation gaps is quite common, and notable examples include the hidden clique problem and the closely related problem of densest submatrix detection (for an overview, see [40]). To achieve the latter case, the optimization algorithms must exploit specific properties of the random instances, as the optimization problem is NP-hard for worst case. A successful example of designing a polynomial-time approximation scheme is given in [36], where the author constructed a greedy algorithm that finds near ground states of certain Gaussian processes on the n -dimensional sphere with high probability. Building on this work, [26] presented an efficient approximation scheme for finding near ground state of the Sherrington-Kirkpatrick model, and the algorithms can also be tailored to find near-optimal solutions for the max-cut problem in dense Erdős-Rényi graphs. It is worth noting that all the efficient approximation algorithms mentioned above rely on a specific property of the underlying stochastic models called *full replica symmetry breaking*. Thus, it is tempting to conjecture that such a property is indicative of ideal performances of certain well-designed efficient algorithms.

Over the past few decades, various frameworks have been proposed to provide insights into the computational hardness of optimization problems with random inputs (for surveys, see [43, 6, 33, 14]). Notably, the overlap-gap property, initially introduced in [17], serves as a geometric barrier that provides solid evidence for the failure of stable algorithms to find solutions beyond a certain

threshold [14]. In the last ten years, several generations of the initial overlap-gap property have been discovered, such as the multi-OGP [34] and the ladder-OGP [39], which establish computational hardness results close to the believed computational thresholds. Recently, a seminal work [21] introduced yet another variant of the overlap-gap property, called the *branching-OGP*, which was used to provide tight and satisfactory hardness results for stable algorithms in mean-field spin glasses. The current work is also significantly inspired by the branching-OGP framework.

Prior work In the paper [12], the authors developed a polynomial-time approximation scheme for the random GAP when the parameter p satisfies $p = n^{-\alpha+o(1)}$ for some constant $\alpha \in (1/2, 1]$ and $p \gg \log n/n$. Despite their sophisticated forms, these algorithms are rooted in a fairly simple idea. Consider a naive greedy graph alignment algorithm, where $\pi(1), \pi(2), \dots, \pi(n)$ are determined sequentially, and for each k , $\pi(k)$ is chosen to maximize $\sum_{i < k} G_{ik} G_{\pi(i)\pi(k)}$. Interestingly, the authors of [12] observed that for certain values of α , this simple algorithm produces near-optimal solutions with high probability. This indicates that the greedy matching algorithm somehow captures the essence of the random GAP problem. Consequently, it is not surprising that all the algorithms presented in [12] and the algorithms that will be discussed in this paper are essentially variants of this simple greedy algorithm.

Acknowledgements The authors would like to give great gratitudes to Jian Ding, for his introduction of this intriguing problem, enlightening comments on conceptual pictures and helpful suggestions for the manuscript. We appreciate Nike Sun a lot for her guidance to the crucial literature [21]. We also thank Brice Huang, Zhangsong Li and Mark Sellke for stimulating discussions.

5.1.2 Proof overview

The proofs of the main results can be divided into four parts, say, informational/computational upper/lower bounds.

The informational upper bounds can be obtained for both the sparse and dense regimes using a simple union bound argument. For the informational lower bound in the dense regime, we observe that the behavior of extremes in the set $\tilde{O}(\pi)_{\pi \in S_n}$ exhibits characteristics similar to Gaussian distributions. Motivated by this, we employ a truncated second-moment method argument to derive the lower bound. It is important to note that a straightforward application of the Paley-Zygmund inequality only yields a vanishing lower bound of probability. However, by combining the use of Talagrand's concentration inequality, we enhance the lower bound to $1 - o(1)$.

In the case where p lies in the sparse regime, surprisingly, we have not discovered any non-constructive proof of the informational lower bound. In fact, Theorem 5.1 is established by proving Theorem 5.3. In other words, we prove the lower bound by constructing specific polynomial-time algorithms that can find solutions which are optimal up to a multiplicative constant arbitrarily close to 1. The most challenging part of the analysis (when $p = n^{-\alpha+o(1)}$ for some $\alpha > 1/2$) has been addressed in [12]. For the remaining parameter ranges in this paper, slight variations of the greedy algorithm mentioned in the last subsection would suffice. The design of the online algorithms \mathcal{A}^* in the dense regime is also inspired by the greedy algorithm. In fact, all the algorithms utilized in this paper can be formulated within a unified framework that draws upon the spirit of the greedy algorithm.

It is believed that the greedy algorithm itself achieves ideal performance in the corresponding regimes, but we still make some modifications to the algorithm for the sake of analysis convenience. The main difficulty in analyzing the performance of such algorithms lies in mitigating the negative impact caused by previous iterations on the current step. We address this issue by employing different approaches in various regimes, tailored to the specific challenges each regime presents.

Finally, we arrive at the hardness result for online algorithms. Our proof combines the resampling technique inspired by [16] with the branching overlap gap property framework presented by [21]. We consider a set of L (which is a large constant chosen in relation to ε) correlated instances $(G_i, \mathbf{G}_i), i = 1, 2, \dots, L$, with a carefully designed ultrametric tree structure of correlation. Roughly speaking, our argument is divided into two parts. Firstly, using the truncated first moment method, we demonstrate that typically it is highly unlikely (with probability less than $\exp(-\Omega(n \log n))$) for certain forbidden structures of permutations $\pi_i \in S_{\beta_c + \varepsilon}(G_i, \mathbf{G}_i), 1 = 1, 2, \dots, L$ to occur. This establishes a strong negative correlation between the desired solutions and the specific structures we consider. Secondly, for each online algorithm \mathcal{A} that finds solutions in $S_{\beta_c + \varepsilon}(G, \mathbf{G})$ with a success probability p_{suc} , we construct such forbidden structures by leveraging \mathcal{A} itself. The probability of successfully constructing these forbidden structures through \mathcal{A} is at least p_{suc}^L , which ensures that the occurrence of these forbidden structures is highly probable under the given conditions. Combining these two parts of the argument leads us to the desired result, demonstrating the hardness of finding solutions beyond the threshold β_c for online algorithms with a probability exceeding $\exp(-\Omega(n \log n))$ under typical instances.

The paper is structured as follows: In Section 5.2, we present the informational results, which

include the upper bound of Theorem 5.1 as well as the corresponding lower bound in the dense regime. In Section 5.3, we construct and analyze various variants of the greedy algorithm. These results complete the proof of Theorem 5.1 and establish the first part (i) of Theorem 5.4. Section 5.4 is dedicated to the hardness result of online algorithms. Here, we provide a detailed analysis that completes the proof of Theorem 5.4, demonstrating the limitations of online algorithms in finding solutions beyond the threshold β_c .

5.2 Proof of informational results

5.2.1 Informational upper bounds

In this short subsection we prove the upper bounds for Theorem 5.1. First we recall the Chernoff bound for Binomial variables.

Proposition 5.6 (Chernoff bound). *For any $N \in \mathbb{N}, P \in (0, 1)$ and $\delta > 0$, let $X \sim \mathbf{B}(N, P)$, it holds that*

$$\mathbb{P}[X \geq (1 + \delta)NP] \leq \exp\left(-NP((1 + \delta)\log(1 + \delta) - \delta)\right), \quad (5.5)$$

and

$$\mathbb{P}[X \leq (1 - \delta)NP] \leq \exp\left(-\frac{\delta^2 NP}{2}\right). \quad (5.6)$$

In particular, we have for any $K \geq 0$,

$$\mathbb{P}[|X - NP| \geq K] \leq 2 \exp\left(-\frac{K^2}{2(NP + K)}\right). \quad (5.7)$$

The proof of (5.5) and (5.6) can be found in, for example, [41, Appendix C, Lemma 11]. Subsequently, (5.7) follows from straightforward algebraic manipulations. The remaining proof involves a straightforward application of the union bound.

Proof of the upper bound for Theorem 5.1. For any fixed permutation $\pi \in S_n$, it is evident that $O(\pi) \sim \mathbf{B}(\binom{n}{2}, p^2)$. We will show that for any fixed $\varepsilon > 0$, the probability of such a binomial variable deviating from its expectation by more than $(1 + \varepsilon)S_{n,p}$ or $(1 + \varepsilon)D_{n,p}$ is much less than $1/n!$. This establishes the desired upper bound through a simple union bound.

When $\log n/n \ll p \ll p_c$, we have $E_{n,p} \ll S_{n,p}$, so by the Chernouff bound,

$$\begin{aligned}
& \mathbb{P} \left[\mathbf{B} \left(\binom{n}{2}, p^2 \right) \geq (1 + \varepsilon) S_{n,p} + E_{n,p} \right] \\
& \stackrel{(5.5)}{\leq} \exp \left(- \left[(E_{n,p} + (1 + \varepsilon) S_{n,p}) \log \left(1 + \frac{(1 + \varepsilon) S_{n,p}}{E_{n,p}} \right) - (1 + \varepsilon) S_{n,p} \right] \right) \\
& \leq \exp \left(- (1 + \varepsilon + o(1)) S_{n,p} \log \frac{S_{n,p}}{E_{n,p}} \right) = \exp \left(- S_{n,p} \cdot (1 + \varepsilon + o(1)) \log \left(\frac{\log n}{np^2} \right) \right) \\
& = \exp \left(- (1 + \varepsilon + o(1)) n \log n \right),
\end{aligned}$$

which is much less than $1/n!$, as desired. When $p_c \ll p \ll 1$, we have $S_{n,p} \ll E_{n,p}$, therefore

$$\mathbb{P} \left[\mathbf{B} \left(\binom{n}{2}, p^2 \right) \geq (1 + \varepsilon) D_{n,p} + E_{n,p} \right] \stackrel{(5.7)}{\leq} 2 \exp \left(- \frac{(1 + \varepsilon)^2 D_{n,p}^2}{2E_{n,p} + 2D_{n,p}} \right),$$

which equals to $\exp \left(- [(1 + \varepsilon)^2 + o(1)] n \log n \right) \ll 1/n!$, as desired. \square

5.2.2 Informational lower bounds for $p \gg p_c$

We denote \mathbf{U} for the set of unordered pairs $\{(i, j) : i, j \in [n], i \neq j\}$. Fix G with edge set $E \subset \mathbf{U}$ and some $\pi \in S_n$, consider the set

$$\text{OL}(G, \pi) = \{(i, j) \in \mathbf{U} : G_{ij} = G_{\pi(i)\pi(j)} = 1\}.$$

then for any $\pi_1, \pi_2 \in S_n$, denote $\pi_{1,2} = \pi_1^{-1} \circ \pi_2$, we claim that it holds

$$\text{Cov}(\tilde{\text{O}}(\pi_1), \tilde{\text{O}}(\pi_2)) = p(1 - p) \times |\text{OL}(G, \pi_{1,2})|.$$

This is because we can express $\text{Cov}(\tilde{\text{O}}(\pi_1), \tilde{\text{O}}(\pi_2)) = \text{Cov}(\text{O}(\pi_1), \text{O}(\pi_2))$ by

$$\sum_{\substack{(i_1, j_1) \in E \\ (i_2, j_2) \in E}} \text{Cov}(\mathbf{G}_{\pi_1(i_1)\pi_1(j_1)}, \mathbf{G}_{\pi_2(i_2)\pi_2(j_2)}) = \sum_{\substack{(i_1, j_1) \in E \\ (i_2, j_2) \in E}} p(1 - p) \mathbf{1}_{(\pi_1(i_1), \pi_1(j_1)) = (\pi_2(i_2), \pi_2(j_2))}$$

and the number of pairs $(i_1, j_1), (i_2, j_2)$ such that $(\pi_1^{-1}(i_1), \pi_1^{-1}(j_1)) = (\pi_2^{-1}(i_2), \pi_2^{-1}(j_2))$ is exactly $|\text{OL}(G, \pi_{1,2})|$.

For each $\pi \in S_n$, denote $F(\pi)$ and $T(\pi)$ for the number of fixed points and transpositions in π , respectively. We note that for $G \sim \mathbf{G}(n, p)$, an unordered pair $(i, j) \in \mathbf{U}$ appears in $\text{OL}(G, \pi)$ with probability p or p^2 , depending on $(i, j) = (\pi(i), \pi(j))$ or not. For the former case to happen, either i, j are both fixed or (i, j) is a transposition. This leads to the following relation

$$\mathbb{E}|\text{OL}(G, \pi)| = \left[\binom{F(\pi)}{2} + T(\pi) \right] (p - p^2) + \binom{n}{2} p^2.$$

The next lemma shows that for $p_c \ll p \ll 1$ and $G \sim \mathbf{G}(n, p)$, the size of $\text{OL}(G, \pi)$ indeed well concentrates around its expectation *uniformly* for all $\pi \in S_n$.

Proposition 5.7. *Denote \mathcal{G} for the event that*

- *The total number of edges in G satisfies*

$$\left| |E(G)| - \binom{n}{2} p \right| \leq \sqrt{n^2 p \log n}. \quad (5.8)$$

- *for any $\pi \in S_n$ it holds*

$$\left| |\text{OL}(G, \pi)| - \mathbb{E} |\text{OL}(G, \pi)| \right| \leq 4\sqrt{n^3 p \log n}. \quad (5.9)$$

Then for $p \gg p_c$, we have $\mathbb{P}[\mathcal{G}] = 1 - o(1)$, where the probability is taken over the random graph $G \sim \mathbf{G}(n, p)$.

Proof. It suffice to prove that both items are true with probability $1 - o(1)$. For the latter one, we have $|E(G)| \sim \mathbf{B}(\binom{n}{2}, p)$, thus by (5.7) we get

$$\mathbb{P} \left[\left| |E| - \binom{n}{2} p \right| \geq \sqrt{n^2 p \log n} \right] \leq 2 \exp \left(- \frac{n^2 p \log n}{n^2 p + 2\sqrt{n^2 p \log n}} \right) = o(1).$$

This shows (5.8) happens with high probability.

For the second item, we recall U is the set of unordered pairs. For any fixed $\pi \in S_n$, let $\Pi : U \rightarrow U$ be defined as $\Pi(i, j) = (\pi(i), \pi(j))$. It is clear that Π is a bijection on U . For a pair $(i, j) \in U$, let $X_{i,j} = G_{i,j} G_{\pi(i), \pi(j)}$, we will use the following observations:

- (i) $X_{i,j}$ is distributed as $\mathbf{B}(1, p)$ if $\Pi(i, j) = (i, j)$, or $\mathbf{B}(1, p^2)$ otherwise.
- (ii) For each $(i, j) \in U$, $X_{i,j}$ is independent with $\{X_{k,l}\}_{(k,l) \in U \setminus \{\Pi(i,j), \Pi^{-1}(i,j)\}}$.

Motivated by this, we may divide U into the disjoint union of four parts $F \sqcup A \sqcup B \sqcup C$, where F is the set of fixed points of Π , and A, B, C are chosen to satisfy that there is no $(i, j) \in U$ such that (i, j) and $\Pi(i, j)$ are in the same set (Such partition can be constructed by a simple iterative procedure).

Denote

$$\sigma(\star) = \sum_{(i,j) \in \star} X_{i,j}, \quad \forall \star \in \{F, A, B, C\},$$

then $\text{OL}(G, \pi) = \sum_{\star \in \{F, A, B, C\}} \sigma(\star)$. In addition, combining the previous observations with the choices of F, A, B, C we made, we see marginally

$$\sigma(F) \sim \mathbf{B}(|F|, p), \text{ and } \sigma(\star) \sim \mathbf{B}(|\star|, p^2), \forall \star \in \{A, B, C\}.$$

Note that $4\sqrt{n^3 p \log n} \geq 2\sqrt{|F| np \log n} + 3\sqrt{2n^3 p^2 \log n}$. Therefore, the probability that $|\text{OL}(G, \pi)|$ deviates from its expectation by at least $4\sqrt{n^3 p \log n}$ is bounded by

$$\begin{aligned} & \mathbb{P} \left[|\sigma(F) - \mathbb{E}\sigma(F)| \geq 2\sqrt{|F| np \log n} \right] + \sum_{\star \in \{A, B, C\}} \mathbb{P} \left[|\sigma(\star) - \mathbb{E}\sigma(\star)| \geq \sqrt{2n^3 p^2 \log n} \right] \\ & \stackrel{(5.7)}{\leq} \exp \left(-\frac{4|F| np \log n}{2|F| p + 4\sqrt{|F| np \log n}} \right) + \sum_{\star \in \{A, B, C\}} \exp \left(-\frac{2n^3 p^2 \log n}{2|\star| p^2 + 2\sqrt{2n^3 p^2 \log n}} \right). \end{aligned}$$

Note that it trivially holds $|\star| \leq n^2/2, \forall \star \in \{A, B, C\}$, so the above expression is much less than $\exp(-n \log n)$. Therefore, the second item is true with high probability by a union bound and the proof is completed. \square

In this section, we will henceforth fix G and assume that \mathcal{G} holds. Notice that $|E(G)| = [1 + o(1)]n^2 p/2$, and we can see from (5.9) that $|\text{OL}(G, \pi)|/|\mathbb{E} \text{OL}(G, \pi)| = 1 + o(1)$ for any $\pi \in S_n$ (since $n^2 p \gg \sqrt{n^3 p \log n}$ when $p \gg p_c$).

Proposition 5.8. *For any graph G which satisfies \mathcal{G} and any constant $\varepsilon > 0$, there exists a constant $\gamma = \gamma(\varepsilon) > 0$, such that*

$$\mathbb{P} \left[\left| \max_{\pi \in S_n} \tilde{\text{O}}(\pi) - \mathbb{E} \max_{\pi \in S_n} \tilde{\text{O}}(\pi) \right| \geq \varepsilon D_{n,p} \right] \leq \exp(-\gamma n \log n),$$

where the probability and expectation is taken over $\mathbf{G} \sim \mathbf{G}(n, p)$.

The result of Proposition 5.8 follows from a standard application of Talagrand's concentration inequality. For completeness, we present the entire argument as below. We start by recalling some definitions and notations.

Definition 5.9. *Let $\Omega = \prod_{i=1}^N \Omega_i$ be a product space. For a function $h : \Omega \rightarrow \mathbb{R}$, we call h to be Lipschitz, if $|h(x) - h(y)| \leq 1$ whenever x, y differ in at most one coordinate. Furthermore, let $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ be an increasing function, we say that h is f -certifiable if, whenever $h(x) \geq s$, there exists $I \subset \{1, \dots, \binom{n}{2}\}$ with $|I| \leq f(s)$ so that all $y \in \Omega$ that agree with x on the coordinates in I have $h(y) \geq s$.*

A special case of Talagrand's concentration inequality can be stated as follow.

Proposition 5.10. *For any product probability measure \mathbb{P} on Ω and any Lipschitz function h that is f -certifiable, let $X = h(\omega)$, we have for all $b \in \mathbb{N}$ and $t > 0$:*

$$\mathbb{P}\left[X \leq b - t\sqrt{f(b)}\right] \mathbb{P}[X \geq b] \leq e^{-t^2/4}.$$

The proof of this powerful inequality can be found in, e.g. [1, Section 7.7]. Now we are ready to give the proof of Proposition 5.8.

Proof. Fix G which ensures \mathcal{G} . For our purposes, we consider $\Omega = \{0, 1\}^{\binom{n}{2}}$ and \mathbb{P} be the product measure of $\binom{n}{2}$ Bernoulli variables with parameter p . Denote

$$X = h(\mathbf{G}) \triangleq \max_{\pi \in S_n} O(\pi),$$

where we view \mathbf{G} as a vector in $\{0, 1\}^{\binom{n}{2}}$. It is trivial to verify that h is Lipschitz.

Moreover, we claim that h defined as above is f -certifiable with $f(s) \triangleq \lceil s \rceil$. Indeed, if $h(x) \geq s$ for some $x = (\mathbf{G}_{ij})_{(i,j) \in \mathbf{U}} \in \Omega$, there exists a permutation π^* and $\lceil s \rceil$ unordered pairs $(i_k, j_k) \in \mathbf{U}$, $1 \leq k \leq \lceil s \rceil$, such that $G_{i_k, j_k} = 1$ and $G_{\pi^*(i_k)\pi^*(j_k)} = 1$ for each k . We can then pick I as $\{(\pi^*(i_k), \pi^*(j_k)), 1 \leq k \leq \lceil s \rceil\}$, then $|I| = \lceil s \rceil$ and for any \mathbf{G}' which contains those edges in I , it holds that $h(\mathbf{G}') \geq O(\pi^*) \geq s$.

As a result, Proposition 5.10 yields that for any $b, t \geq 0$,

$$\mathbb{P}[X \leq b - t\sqrt{\lceil b \rceil}] \mathbb{P}[X \geq b] \leq e^{-t^2/4}. \quad (5.10)$$

Taking $b = \text{Median}(X)$ in (5.10) and by an easy transformation, we get for any $t \geq 0$,

$$\mathbb{P}(X - \text{Median}(X) \leq -t) \leq 2 \exp\left(-\frac{t^2}{4\lceil \text{Median}(X) \rceil}\right).$$

In addition, for any $r \geq 0$, taking $b = \text{Median}(X) + r$, $t = r/\sqrt{b}$ in (5.10) yields:

$$\begin{aligned} \mathbb{P}[X - \text{Median}(X) \geq r] &\leq 2 \exp\left(-\frac{r^2}{4(\lceil \text{Median}(X) \rceil + r)}\right) \\ &\leq 2 \exp\left(-\frac{r^2}{8\lceil \text{Median}(X) \rceil}\right) + 2 \exp\left(-\frac{r^2}{8\lceil r \rceil}\right). \end{aligned}$$

From a similar argument as presented in Subsection 5.2.1, we see conditioned on G satisfies \mathcal{G} , $\text{Median}(X)$ is of order $\mathbb{E}O(\pi) \sim n^2 p^2 / 2$, in particular, $\lceil \text{Median}(X) \rceil \leq n^2 p^2$. We now claim that

$|\mathbb{E}[X] - \text{Median}(X)| = o(D_{n,p})$. Indeed using the aforementioned tail estimations we have

$$\begin{aligned}
& |\mathbb{E}[X] - \text{Median}(X)| \leq \mathbb{E}[|X - \text{Median}(X)|] \\
&= \int_0^{+\infty} \mathbb{P}(|X - \text{Median}(X)| \geq t) dt \\
&\leq \int_0^{+\infty} \left[2 \exp\left(-\frac{t^2}{4n^2p^2}\right) + 2 \exp\left(-\frac{t^2}{8n^2p^2}\right) + 2 \exp\left(-\frac{t^2}{8\lceil t \rceil}\right) \right] dt \\
&= O(np) = o(D_{n,p}).
\end{aligned} \tag{5.11}$$

As a result, we see for n large enough,

$$\begin{aligned}
& \mathbb{P}[|X - \mathbb{E}[X]| \geq \varepsilon D_{n,p}] \leq \mathbb{P}[|X - \text{Median}(X)| \geq \varepsilon D_{n,p}/2] \\
&\leq 2 \exp\left(-\frac{\varepsilon^2 D_{n,p}^2}{16n^2p^2}\right) + 2 \exp\left(-\frac{\varepsilon^2 D_{n,p}^2}{32n^2p^2}\right) + 2 \exp\left(-\frac{\varepsilon D_{n,p}}{16}\right) \\
&= \exp(-\Omega(n \log n)),
\end{aligned}$$

as desired. \square

We assume $p_c \ll p \ll 1$ in this subsection. Fix G such that \mathcal{G} holds, for any $\varepsilon > 0$, define X_ε as the number of $\pi \in S_n$ such that $\tilde{O}(\pi) \geq (1 - \varepsilon)D_{n,p}$. We shall prove the following estimate.

Proposition 5.11. *Conditioned on any realization of G that satisfies the event \mathcal{G} , for any constant $\varepsilon \in (0, 1)$, it holds that*

$$\mathbb{E}X_\varepsilon^2 = \exp(o(n \log n))(\mathbb{E}X_\varepsilon)^2.$$

Assume Proposition 5.11, we may prove the lower bound of Theorem 5.1 by a combination of the Paley-Zygmund inequality and the concentration result stated as before.

Corollary 5.12. *For any $\varepsilon > 0$ we have as $n \rightarrow \infty$,*

$$\mathbb{P}\left[\max_{\pi \in S_n} \tilde{O}(\pi) \geq (1 - \varepsilon)D_{n,p}\right] = 1 - o(1).$$

Proof. Since $\mathbb{P}[\mathcal{G}] = 1 - o(1)$ we may condition on the realization of G and assume \mathcal{G} holds. From Proposition 5.11 and the Paley-Zygmund inequality, we have

$$\mathbb{P}\left[\max_{\pi \in S_n} \tilde{O}_\pi \geq (1 - \varepsilon/2)D_{n,p}\right] = \mathbb{P}[X_{\varepsilon/2} > 0] \geq \frac{(\mathbb{E}X_{\varepsilon/2})^2}{\mathbb{E}X_{\varepsilon/2}^2} = \exp(-o(n \log n)).$$

Combining with Proposition 5.8, this implies $\mathbb{E}\max_{\pi \in S_n} \tilde{O}_\pi \geq (1 - 3\varepsilon/4)D_{n,p}$. Then by Applying the result of Proposition 5.8 again we reach the conclusion. \square

Now we turn to the proof of Proposition 5.11. We start by the following two point estimation.

Proposition 5.13. *For $p_c \ll p \ll 1$ and any G satisfying \mathcal{G} , we have for any $\varepsilon \in (0, 1)$ and any $\pi_1, \pi_2 \in S_n$, write $\pi_{1,2} = \pi_1^{-1} \circ \pi_2$, then*

$$\begin{aligned} & \mathbb{P}[\tilde{O}(\pi_1) \geq (1 - \varepsilon)D_{n,p}, \tilde{O}(\pi_2) \geq (1 - \varepsilon)D_{n,p}] \\ & \leq \exp\left(-\frac{2(1 - \varepsilon)^2 n \log n}{1 + (F(\pi_{1,2})/n)^2} + o(n \log n)\right). \end{aligned} \quad (5.12)$$

Proof. Recall that U is the set of unordered pairs and $OL(G, \pi_{1,2})$ is the subset of U that

$$\{(i, j) \in U : G_{i,j} = G_{\pi_{1,2}(i), \pi_{1,2}(j)} = 1\} = \{(i, j) \in U : G_{\pi_1^{-1}(i), \pi_1^{-1}(j)} = G_{\pi_2^{-1}(i), \pi_2^{-1}(j)} = 1\}.$$

We write $L = |OL(G, \pi_{1,2})|$ for simplicity, and we denote

$$\begin{aligned} S_0 &= \sum_{(i,j) \in OL(G, \pi_{1,2})} G_{\pi_1^{-1}(i), \pi_1^{-1}(j)} G_{i,j}, \\ S_1 &= \sum_{(i,j) \in U \setminus OL(G, \pi_{1,2})} G_{\pi_1^{-1}(i), \pi_1^{-1}(j)} G_{i,j}, \\ S_2 &= \sum_{(i,j) \in U \setminus OL(G, \pi_{1,2})} G_{\pi_2^{-1}(i), \pi_2^{-1}(j)} G_{i,j}. \end{aligned}$$

It is clear that $S_0 \sim \mathbf{B}(L, p)$ and $S_1, S_2 \sim \mathbf{B}(E - L, p)$ are independent binomial variables, and

$$\begin{aligned} O(\pi_1) &= \sum_{(i,j) \in E} G_{\pi_1^{-1}(i), \pi_1^{-1}(j)} G_{i,j} = S_0 + S_1, \\ O(\pi_2) &= \sum_{(i,j) \in E} G_{\pi_2^{-1}(i), \pi_2^{-1}(j)} G_{i,j} = S_0 + S_2 \end{aligned}$$

Denote $M = \lceil (1 - \varepsilon)D_{n,p} \rceil$, then we have $\mathbb{P}[\tilde{O}(\pi_1) \geq M, \tilde{O}(\pi_2) \geq M]$ equals to

$$\begin{aligned} & \mathbb{P}[S_0 + S_1 \geq Ep + M, S_0 + S_2 \geq Ep + M] \\ &= \sum_k \mathbb{P}[S_0 = Lp + k] (\mathbb{P}[S_1 \geq (E - L)p + M - K])^2 \\ &\leq \mathbb{P}[S_0 \geq Lp + M] + (\mathbb{P}[S_1 \geq (E - L)p + M])^2 \\ &\quad + (M + 1) \max_{0 \leq k \leq M} \mathbb{P}[S_0 \geq Lp + k] (\mathbb{P}[S_1 \geq (E - L)p + M - k])^2. \end{aligned} \quad (5.13)$$

We claim that each term in (5.13) is bounded by the right hand side of (5.12). Note that by the condition (5.9) in \mathcal{G} , we have $L/E = (F(\pi_{1,2})/n)^2 + o(1)$. By applying (5.7), we see first term of

the expression is bounded by

$$\exp\left(-\frac{(1-\varepsilon)^2 D_{n,p}^2}{2(Lp + D_{n,p})}\right) \leq \begin{cases} \exp(-2(1-\varepsilon)^2 n \log n + o(n \log n)), & \text{if } F(\pi_{1,2})/n \leq 1/2, \\ \exp\left(-\frac{(1-\varepsilon)^2 n \log n}{(F(\pi_{1,2})/n)^2} + o(n \log n)\right), & \text{if } F(\pi_{1,2})/n > 1/2, \end{cases}$$

which is always bounded by $\exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1+(F(\pi_{1,2})/n)^2} + o(n \log n)\right)$, as desired. Similar arguments suggest that this is true for the second term in (5.13). For the last term, applying (5.7) again we see it is bounded by $M + 1 = \exp(o(n \log n))$ times

$$\begin{aligned} \max_{1 \leq k \leq M} \exp\left(-\frac{k^2}{2(Lp + k)} - \frac{(M - k)^2}{(E - L)p + M - k}\right) &\leq \max_{0 \leq k \leq M} \exp\left(-\frac{M^2}{Ep + Lp + M + k}\right) \\ &= \exp\left(-\frac{2(1-\varepsilon)^2 n^3 p^2 \log n}{Ep + Lp + 2M}\right) = \exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1 + L/E + o(1)}\right) \\ &= \exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1 + (F(\pi_{1,2})/n)^2} + o(n \log n)\right), \end{aligned}$$

completing the proof of the claim. Therefore, (5.12) follows readily. \square

Proof of Proposition 5.11. By definition we have

$$\mathbb{E}X_\varepsilon^2 = \sum_{\pi_1, \pi_2 \in S_n} \mathbb{P}[\tilde{O}(\pi_1) \geq (1-\varepsilon)D_{n,p}, \tilde{O}(\pi_2) \geq (1-\varepsilon)D_{n,p}].$$

From Proposition 5.13 we get this is bounded by

$$\begin{aligned} &\exp(o(n \log n)) \times \sum_{\pi_1, \pi_2 \in S_n} \exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1 + (F(\pi_{1,2})/n)^2}\right) \\ &= \exp([1 + o(1)]n \log n) \sum_{\pi \in S_n} \exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1 + (F(\pi)/n)^2}\right) \\ &= \exp([1 + o(1)]n \log n) \times \sum_{k=1}^n \exp\left(-\frac{2(1-\varepsilon)^2 n \log n}{1 + (k/n)^2}\right) \times |\pi \in S_n : F(\pi) = k| \\ &\leq \exp([2 + o(1)]n \log n) \times \sum_{k=1}^n \exp\left(-\left[\frac{2(1-\varepsilon)^2}{1 + (k/n)^2} + k/n\right]n \log n\right) \\ &\leq \exp([2 - 2(1-\varepsilon)^2 + o(1)]n \log n), \end{aligned}$$

where in the first inequality we used a well-known fact that $|\pi \in S_n : F(\pi) = k|$ is bounded by $\exp((n - k) \log n + o(n \log n))$ and the second inequality follows from the observation that for any $\gamma \in [0, 1]$, it holds

$$\gamma[1 - 2(1-\varepsilon)^2\gamma + \gamma^2] \geq 0 \iff \frac{2(1-\varepsilon)^2}{1 + \gamma^2} + \gamma \geq 2(1-\varepsilon)^2.$$

Finally, we conclude the proof of Proposition 5.11 by noting that from Lemma 5.25,

$$\mathbb{E}X_\varepsilon = n! \times \mathbb{P}[\tilde{O}(\pi) \geq (1 - \varepsilon)D_{n,p}] \geq \exp([1 - (1 - \varepsilon)^2 + o(1)]n \log n). \quad \square$$

5.3 Algorithmic lower bounds via greedy matching

In this section, we construct and analyze matching algorithms that have the desired properties stated in Theorem 5.3 and (i) in Theorem 5.4. Note that for the case when $p = n^{-\alpha+o(1)}$ with some constant $1/2 < \alpha < 1$, a polynomial-time approximation scheme has been presented in [12], so this paper addresses no attention to this (super-sparse) regime. For the remaining regimes (i.e. $p = n^{-1/2+o(1)}$ and $p \ll p_c$ or $p_c \ll p \leq 1/(\log n)^3$), we will apply algorithms from a simplified and unified framework.

5.3.1 The algorithm framework and proof outlines

We present the following variant of the greedy matching algorithm stated in Section 1.3.

Algorithm 2 Greedy Matching Algorithm

- 1: **Input:** an losing parameter $\eta > 0$, the targets F_s in each step $\eta n \leq s \leq (1 - \eta)n$, and the adjacency matrices $\{G_{ij}\}, \{G_{ij}\}$ for G, G respectively.
- 2: **Initialize:** $\pi^*(i) = i$ for $1 \leq i \leq \eta n$, $R_{\lfloor \eta n \rfloor} = \{\lfloor \eta n \rfloor + 1, \dots, n\}$.
- 3: **for** $\lfloor \eta n \rfloor + 1 \leq s \leq n$ **do**
- 4: Set $I_s = 0$ and denote $R_{s-1} = \{j_1, \dots, j_t\}, j_1 < \dots < j_t$.
- 5: **for** $k = 1, \dots, t$ **do**
- 6: **if** $\sum_{i: i < s} G_{i,s} G_{\pi^*(i), j_k} \geq F_s$ **then**
- 7: Set $I_s = 1, \pi^*(s) = j_k, R_s = R_{s-1} \setminus \{j_k\}$ and break the **for** cycle. {Once a vertex that matches up to F_s edges is found, we adapt it and stop searching.}
- 8: **else if** $k = t$ **then**
- 9: Set $\pi^*(i) = j_1, R_s = R_{s-1} \setminus \{j_1\}$. {If no such vertex exists, we do the matching in a somewhat arbitrary way.}
- 10: **end if**
- 11: **end for**
- 12: **end for**
- 13: **Output:** π^* .

We will choose appropriate parameter η (sufficiently small) and targets $F_s, \eta n \leq s \leq (1 - \eta)n$ for p in various regimes, and show that the output π^* of the algorithm above satisfies the desired property with high probability.

We begin with some notations. Write $S(\eta) = \{\lfloor \eta n \rfloor + 1, \dots, \lfloor (1 - \eta)n \rfloor\}$ for simplicity. For $s \in S(\eta)$, we let \mathcal{F}_{s-1} denote the σ -field generated by $\pi^*(i), i < s$ and the matched parts of graphs $\{G_{i,j}, G_{\pi^*(i), \pi^*(j)}, i < j < s\}$ before the s -th step. In addition, we let $\mathcal{N}_{s-1} \subset \mathcal{F}_{s-1}$ be the σ -field generated by $\{G_{i,j} : i < j < s\}$ and $N_k = \{j < k : G_{j,k} = 1\}$ for $k \leq s$. For each $r \in R_{s-1}$, denote

$$E_{r,s} = \sum_{j < s} G_{j,s} G_{\pi^*(j),r} = \sum_{j \in N_s} G_{\pi^*(j),r}.$$

Furthermore, conditioned on any realization of \mathcal{F}_{s-1} , we denote E_s for the set of variables $G_{i,j}, i \in R_{s-1}, j \in R_{s-1}^c$ and let $\tilde{\mathcal{N}}_s$ be the σ -field generated by variables in E_s . We remark that despite the similarity between the notations \mathcal{N}_s and $\tilde{\mathcal{N}}_s$, they represent totally different information in the two graphs G and G .

Before diving into details, we give a sketch to our approach first. Heuristically, we would like to show that most steps in the matching algorithm “succeeds”, i.e. $I_s = 1$ for most indices s . In order to show this, we prove that under specific choices of η and $F_s, s \in S(\eta)$, it holds for any s that $\mathbb{P}[I_s = 1] = 1 - o(1)$. we will show that for each $s \in S(\eta)$, there exists some appropriate good event \mathcal{G}_{s-1}^1 which is measurable with respect to \mathcal{F}_{s-1} , such that $\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1)$, and for any realization of \mathcal{F}_{s-1} that satisfies \mathcal{G}_{s-1}^1 , it always holds that $\mathbb{P}[I_s = 1 \mid \mathcal{F}_{s-1}] = 1 - o(1)$. Once this is true, by the iterated expectation formula we have

$$\mathbb{P}[I_s = 1] = \mathbb{E}[\mathbb{P}[I_s = 1 \mid \mathcal{F}_{s-1}]] \geq [1 - o(1)]\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1),$$

as desired.

Note that conditioned on the realization of \mathcal{F}_{s-1} , the event $\{I_s = 1\}$ is equivalent to there exists some $i \in R_{s-1}$ such that $E_{i,s} \geq F_s$. Thus we denote

$$X_s = \sum_{r \in R_{s-1}} \mathbf{1}_{\{E_{r,s} \geq F_s\}},$$

then $\mathbb{P}[I_s = 1 \mid \mathcal{F}_{s-1}] = \mathbb{P}[X_s > 0 \mid \mathcal{F}_{s-1}]$. We will use the truncated second moment method to bound such probability from below. More precisely, we will introduce yet another good event \mathcal{G}_s^2 , which may differ for p in various regimes, but is measurable with respect to $\sigma(\mathcal{F}_{s-1} \cup \mathcal{N}_s)$ or

$\sigma(\mathcal{F}_{s-1} \cup \tilde{\mathcal{N}}_s)$. The choice of such good event \mathcal{G}_s^2 will satisfy that under any realization of \mathcal{F}_{s-1} the satisfies \mathcal{G}_{s-1}^1 , $\mathbb{P}[\mathcal{G}_s^2 \mid \mathcal{F}_{s-1}] = 1 - o(1)$, and moreover, under any realization of \mathcal{N}_s (or $\tilde{\mathcal{N}}_s$) that satisfies \mathcal{G}_s^2 it holds that

$$\mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1}, \mathcal{N}_s] = [1 + o(1)] (\mathbb{E}[X_s \mid \mathcal{F}_{s-1}, \mathcal{N}_s])^2, \quad (5.14)$$

or the same holds for \mathcal{N}_s replaced by $\tilde{\mathcal{N}}_s$. Once this is true, we see for any realization \mathcal{F}_{s-1} that satisfies \mathcal{G}_{s-1}^2 , we have

$$\begin{aligned} \mathbb{P}[X_s > 0 \mid \mathcal{F}_{s-1}] &\geq \mathbb{E}[\mathbf{1}_{\mathcal{G}_s^2} \mathbb{P}[X_s > 0 \mid \mathcal{F}_{s-1}, \mathcal{N}_s]] \\ &\geq \mathbb{E}\left[\mathbf{1}_{\mathcal{G}_s^2} \frac{(\mathbb{E}[X_s \mid \mathcal{F}_{s-1}, \mathcal{N}_s])^2}{\mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1}, \mathcal{N}_s]} \mid \mathcal{F}_{s-1}\right] = [1 - o(1)] \mathbb{P}[\mathcal{G}_s^2 \mid \mathcal{F}_{s-1}] = 1 - o(1), \end{aligned}$$

(or the same holds for \mathcal{N}_s replaced by $\tilde{\mathcal{N}}_s$) where in the second second inequality we applied the Paley-Zygmund inequality.

Here in this subsection we also give a first glance at how we will prove such first and second moments estimates as in (5.14). Fix an index $s \in S(\eta)$ and a realization of \mathcal{F}_{s-1} , recall the definition of \mathbf{E}_s , then clearly the event $\{I_s = 1\} = \{X_s > 0\}$ only depends on the variables in \mathbf{E}_s . Denote $\hat{\mathbb{P}}_s$ for the joint law of variables in \mathbf{E}_s (which is a product measure of $\mathbf{B}(1, p)$ variables). According to the rule of our algorithm, conditioned on the \mathcal{F}_{s-1} , it is equivalent to that

$$\sum_{i < k} G_{i,k} \mathbf{G}_{\pi^*(i),j} < F_k, \text{ for any } j \in \mathbf{R}_{s-1} \text{ and } \lfloor \eta n \rfloor < k < s \text{ s.t. } \pi^*(k) > j, \quad (5.15)$$

and $\sum_{i < k} G_{i,k} \mathbf{G}_{\pi^*(i),\pi^*(k)} \geq F_k, \forall \lfloor \eta n \rfloor < k < s$. Denote \mathcal{D}_{s-1} for the event in (5.15), then clearly \mathcal{D}_s is decreasing and we note that $\hat{\mathbb{P}}_s[\cdot \mid \mathcal{F}_{s-1}] = \hat{\mathbb{P}}_s[\cdot \mid \mathcal{D}_{s-1}]$ since the latter part of conditioning is independent with \mathbf{E}_s . We will show that under the good event \mathcal{G}_{s-1}^1 , the decreasing event \mathcal{D}_s does not “tilt” $\hat{\mathbb{P}}_s$ too much in the sense that the (truncated versions of) first and second moments of X_s under $\hat{\mathbb{P}}_s[\cdot \mid \mathcal{D}_s]$ is almost same as the corresponding moments under $\hat{\mathbb{P}}$, and this will yield (5.14).

Now assume $\mathbb{P}[I_s = 1] = 1 - o(1)$ holds for any $s \in S(\eta)$. By Markov Inequality, we see

$$\mathbb{P}\left[\sum_{s \in S(\eta)} \mathbf{1}_{I_s=0} \geq \eta n\right] \leq \frac{\sum_{s \in S(\eta)} \mathbb{P}[I_s = 0]}{\eta n} = o(1). \quad (5.16)$$

Denote \mathcal{S} for the event that there are at least $(1 - 3\eta)n$ indices $s \in S(\eta)$ with $I_s = 1$, then $\mathbb{P}[\mathcal{S}] = 1 - o(1)$. We now fix an arbitrary constant $\eta > 0$ and deal with the sparse regime and the dense regime separately.

Sparse regime For $p = n^{-1/2+o(1)}$ and $p \ll p_c$, we let

$$F_s = \frac{(1-\eta) \log n}{\log(\log n / np^2)}, \quad \forall s \in S(\eta). \quad (5.17)$$

We will show in Subsection 3.2 that for any constant $\eta > 0$ and such choices of F_s , we have $\mathbb{P}[\mathcal{S}] = 1 - o(1)$ as above.

We write the increment of the target quantity $O(\pi^*)$ in step s by

$$\mathcal{O}_s = \sum_{j < s} G_{j,s} \mathbf{G}_{\pi^*(j), \pi^*(s)}. \quad (5.18)$$

Given that \mathcal{S} happens, it follows

$$O(\pi^*) = \sum_{s \in S(\eta)} \mathcal{O}_s \geq \sum_{s \in S(\eta)} F_s \mathbf{1}_{I_s=1} \geq (1-3\eta)n \cdot \frac{(1-\eta) \log n}{\log(\log n / np^2)} \geq (1-4\eta)S_{n,p}.$$

This shows $\mathbb{P}[O(\pi^*) \geq (1-4\eta)S_{n,p}] = 1 - o(1)$ and thus proves Theorem 5.3 since η can be picked arbitrarily small (recall that $E_{n,p} \ll S_{n,p}$ for p in the sparse regime).

dense regime When $p_c \ll p \ll 1/(\log n)^3$, we let

$$F_s = sp^2 + \sqrt{2(1-\varepsilon_n)sp^2 \log n}, \quad \forall s \in S(\eta), \quad (5.19)$$

here $0 < \varepsilon_n \leq \eta$ is a small factor which may depending on p . We will be able to show that under such choices, $\mathbb{P}[\mathcal{S}] = 1 - o(1)$ still holds (see Subsection 3.3). In addition, we need the following proposition to show that the rest steps does not contribute too less to $O(\pi^*)$. Recall the definition of \mathcal{O}_s , $1 \leq s \leq n$ as in (5.18).

Proposition 5.14. *Denote \mathcal{C} for the event that the following items hold:*

- $\sum_{s \leq \lfloor \eta n \rfloor} \mathcal{O}_s = \sum_{i < j \leq \lfloor \eta n \rfloor} G_{i,j} \mathbf{G}_{i,j} \geq \sum_{s \leq \lfloor \eta n \rfloor} sp^2 - \eta D_{n,p}.$
- For each $\lfloor \eta n \rfloor + 1 \leq s \leq n$, $\mathcal{O}_s \geq sp^2 - \sqrt{10np^2 \log n}.$

Then $\mathbb{P}[\mathcal{C}] = 1 - o(1).$

The proof is not difficult, and we leave it in the appendix. Under $\mathcal{S} \cap \mathcal{C}$ we have

$$\begin{aligned}
O(\pi^*) &= \sum_{s \leq \lfloor \eta n \rfloor} \mathcal{O}_s + \sum_{\lfloor \eta n \rfloor + 1 \leq s \leq n} \mathcal{O}_s \\
&\geq \sum_{s \leq \lfloor \eta n \rfloor} \mathcal{O}_s + \sum_{\lfloor \eta n \rfloor + 1 \leq s \leq n} (sp^2 + \sqrt{2(1 - \varepsilon_n)sp^2 \log n} \mathbf{1}_{I_s=1} - \sqrt{10np^2 \log n} \mathbf{1}_{I_s=0}) \\
&\geq \sum_{s \leq n} sp^2 - \eta D_{n,p} + \sqrt{(1 - \varepsilon_n) \cdot 8/9 \cdot n^3 p^2 \log n} - 2\eta n \cdot \sqrt{20np^2 \log n} \\
&\geq \binom{n}{2} p^2 + (\beta_c - 20\eta) D_{n,p},
\end{aligned}$$

where in the third line we used the fact that

$$\sum_{\lfloor \eta n \rfloor + 1 \leq s \leq n} \sqrt{2s} \geq \int_{\lfloor \eta n \rfloor}^n \sqrt{2x} \, dx \geq (\sqrt{8/9} - \eta)n\sqrt{n}.$$

This shows that $\mathbb{P}[\tilde{O}(\pi^*) \geq (\beta_c - 20\eta) D_{n,p}] \geq \mathbb{P}[\mathcal{S} \cap \mathcal{C}] = 1 - o(1)$, thus proves (i) in Theorem 5.4.

5.3.2 Analysis for sparse regime

In this section we assume $p = n^{-1/2+o(1)}$ and $p \ll p_c$. Fix a small constant $\eta \in (0, 0.01)$ and any $s \in \mathcal{S}(\eta)$. Let F_s defined as in (5.17), our goal is to show $\mathbb{P}[I_s = 1] = 1 - o(1)$ under such assumptions. Following the approach suggested in subsection 3.1, we begin with defining the good events \mathcal{G}_{s-1}^1 and \mathcal{G}_s^2 .

Definition 5.15 (Good events in sparse regime). *Denote \mathcal{G}_{s-1}^1 for the event that*

$$(1 - \eta)rp \leq |N_r| \leq (1 + \eta)rp, \forall \lfloor \eta n \rfloor + 1 \leq r \leq s - 1. \quad (5.20)$$

In addition, we let \mathcal{G}_s^2 be the event that $(1 - \eta)sp \leq N_s \leq (1 + \eta)sp$ and

$$|N_r \cap N_s| \leq (\log n)^3, \forall \lfloor \eta n \rfloor + 1 \leq r \leq s - 1. \quad (5.21)$$

Note that \mathcal{G}_{s-1}^1 is measurable with respect to \mathcal{N}_{s-1} (thus also \mathcal{F}_{s-1}) and \mathcal{G}_s^2 is measurable with respect to \mathcal{N}_s . We claim that both events happens with high probability.

Lemma 5.16. *It holds that $\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1)$, and for any realization of \mathcal{F}_{s-1} that satisfying \mathcal{G}_{s-1}^1 , one has $\mathbb{P}[\mathcal{G}_s^2 \mid \mathcal{F}_{s-1}] = 1 - o(1)$.*

Proof. Note that for each r , $|N_i|$ distributed as a binomial variable $\mathbf{B}(r, p)$, and when conditioned on \mathcal{F}_{s-1} , the distribution of $|N_r \cap N_s|$ is given by $\mathbf{B}(|N_r|, p)$, which is stochastically dominated by $\mathbf{B}((1+\eta)rp, p)$ when \mathcal{G}_{s-1}^1 holds. We see from the Chernoff bound (5.7) that for any $\eta n \leq r \leq n$,

$$\mathbb{P}[\mathbf{B}(r, p) \geq (1+\eta)rp] \leq \exp\left(-\frac{\eta^2 rp}{2(1+\eta)}\right) = o(1/n),$$

and (recall that $np^2 \ll \log n$)

$$\mathbb{P}[\mathbf{B}((1+\eta)rp, p) \geq (\log n)^3] \leq \exp\left(-\frac{(\log n)^6}{2((1+\eta)rp^2 + (\log n)^3)}\right) = o(1/n),$$

Thus the lemma follows from a simple union bound \square

Now under these good events, we are able to perform the truncated second moment method. Recall the probability measure $\widehat{\mathbb{P}}_s$ defined as in the previous subsection, which is the product measure of edges in \mathbf{E}_s . We denote the expectation with respect to $\widehat{\mathbb{P}}_s$ by $\widehat{\mathbb{E}}_s$. Formally, we will show the estimates given in the proposition below, which yields (5.14) and thus proves $\mathbb{P}[I_s = 1] = 1 - o(1)$ from the previous argument.

Proposition 5.17. *For each realization \mathcal{F}_{s-1} satisfying \mathcal{G}_{s-1}^1 and \mathcal{N}_s satisfying \mathcal{G}_s^2 , we have,*

$$\mathbb{E}[X_s | \mathcal{F}_{s-1}, \mathcal{N}_s] \geq [1 - o(1)] \widehat{\mathbb{E}}_s[X_s] \rightarrow \infty, \quad (5.22)$$

and

$$\mathbb{E}[X_s^2 | \mathcal{F}_{s-1}, \mathcal{N}_s] \leq [1 + o(1)] (\widehat{\mathbb{E}}_s[X_s])^2. \quad (5.23)$$

We remark that it is straightforward to check $\widehat{\mathbb{E}}_s[X_s] \rightarrow \infty$ (see e.g. Lemma ??). In addition, the second moment estimate (5.23) is fairly easy, since it follows essentially from FKG inequality. By definition we have

$$\begin{aligned} \mathbb{E}[X_s^2 | \mathcal{F}_{s-1}, \mathcal{N}_s] &= \sum_{i,j \in \mathbf{R}_{s-1}} \widehat{\mathbb{P}}_s[E_{i,s} \geq F_s, E_{j,s} \geq F_s | \mathcal{D}_{s-1}] \\ &\leq \sum_{i,j \in \mathbf{R}_{s-1}} \widehat{\mathbb{P}}_s[E_{i,s} \geq F_s, E_{j,s} \geq F_s] \\ &= \sum_{i \neq j \in \mathbf{R}_s} \widehat{\mathbb{P}}_s[E_{i,s} \geq F_s] \widehat{\mathbb{P}}_s[E_{j,s} \geq F_s] + \widehat{\mathbb{E}}_s[X_s] \\ &\leq (\widehat{\mathbb{E}}_s[X_s])^2 + \widehat{\mathbb{E}}_s[X_s] = [1 + o(1)] (\widehat{\mathbb{E}}_s[X_s])^2, \end{aligned}$$

as desired. Where in the proof, we first used FKG inequality since \mathcal{D}_{s-1} is decreasing and $\{E_{i,s} \geq F_s, E_{j,s} \geq F_s\}$ is increasing, and subsequently we used independence.

The remaining of this section is devoted for showing (5.22). We fix realizations of \mathcal{F}_{s-1} that satisfies \mathcal{G}_{s-1}^1 and \mathcal{N}_s that satisfies \mathcal{G}_s^2 . Note that given these realizations, the events $\{E_{r,k} < F_k\}, r \in \mathbf{R}_{s-1}, k \leq s$ are all measurable with respect to \mathbf{E}_s . For each index $r \in \mathbf{R}_{s-1}$, we define $\mathbf{A}_{r,s}$ for the set of indices $k < s$ such that $\{E_{r,k} < F_k\}$ has been verified in the k -th step of iteration (i.e. $I_k = 0$ or $I_k = 1$ and $\pi^*(k) > r$). It is easy to see

$$\mathcal{D}_{s-1} = \bigcap_{r \in \mathbf{R}_{s-1}} \bigcap_{k \in \mathbf{A}_{r,s}} \{E_{r,k} < F_k\} \stackrel{\Delta}{=} \bigcap_{r \in \mathbf{R}_{s-1}} \mathcal{A}_{r,s},$$

where the events $\mathcal{A}_{r,s} = \bigcap_{k \in \mathbf{A}_{r,s}} \{E_{r,k} < F_k\}$ are conditionally independent whenever \mathcal{F}_{s-1} and \mathcal{N}_s are given, since $\mathcal{A}_{r,s}$ is measurable with respect to $\{\mathbf{G}_{\pi^*(i),r} : i < s\}$.

From the above observation we see for each $r \in \mathbf{R}_{s-1}$, the event $\{E_{r,s} \geq F_s\}$ is independent with $\bigcap_{r' \in \mathbf{R}_{s-1} \setminus \{r\}} \mathcal{A}_{r',s}$, so it holds

$$\begin{aligned} \widehat{\mathbb{P}}_s[E_{r,s} \geq F_s \mid \mathcal{D}_{s-1}] &= \widehat{\mathbb{P}}_s[E_{r,s} \geq F_s \mid \mathcal{A}_{r,s}] \geq \frac{\widehat{\mathbb{P}}_s[\{F_s \leq E_{r,s} \leq 6F_s\} \cap \mathcal{A}_{r,s}]}{\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}]} \\ &= \widehat{\mathbb{P}}_s[F_s \leq E_{r,s} \leq 6F_s] \times \frac{\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s} \mid F_s \leq E_{r,s} \leq 6F_s]}{\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}]}. \end{aligned}$$

It can be shown that $\widehat{\mathbb{P}}_s[F_s \leq E_{r,s} \leq 6F_s] = (1 - o(1))\widehat{\mathbb{P}}_s[E_{i,s} \geq F_s]$ (see Lemma 5.24). Therefore, in order to obtain (5.22), it suffices to show,

$$\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s} \mid F_s \leq E_{r,s} \leq 6F_s] \geq [1 - o(1)]\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}]. \quad (5.24)$$

To obtain (5.24), we define for each $r \in \mathbf{R}_{s-1}$ and $k \in \mathbb{N}$ the random set of indices

$$\mathbf{R}_s^\ell := \{k \in \mathbf{A}_{r,s} : \sum_{j \in \mathbf{N}_r \cap \mathbf{N}_s} \mathbf{G}_{k,\pi^*(j)} = \ell\},$$

namely, the set of unmatched vertices in \mathbf{G} that connect exactly ℓ edges to the set $\pi^*(\mathbf{N}_r) \cap \pi^*(\mathbf{N}_s)$. Denote \mathcal{R}_s for the collection of $\{\mathbf{R}_s \mid \mathbf{R}_s = (\mathbf{R}_s^k, k \in \mathbb{N})\}$. We call a collection \mathcal{R}_s to be typical, if $|\mathbf{R}_s^1| + |\mathbf{R}_s^2| < n^{0.6}$ and $\mathbf{R}_s^k = \emptyset, \forall k \geq 3$. The following lemma justifies such calling.

Lemma 5.18. *Whenever \mathcal{F}_{s-1} satisfies \mathcal{G}_{s-1}^1 and \mathcal{N}_s satisfies \mathcal{G}_s^2 , then*

$$\widehat{\mathbb{P}}_s[\mathcal{R}_s \text{ is typical} \mid F_s \leq E_{r,s} \leq 6F_s] \geq 1 - o(1). \quad (5.25)$$

Proof. Fix any integer $K \in [F_s, 6F_s]$, we note that conditioned on the event that $E_{r,s} = K$, the set of indices $i \in N_s$ such that $G_{r,\pi^*}(i) = 1$ is a uniform subset of N_s with K elements. As a result, for each k , the quantity $\sum_{j \in N_k \cap N_s} G_{r,\pi^*}(j)$ distributes like the size of intersection of $N_k \cap N_s$ with a uniform K -subset of N_s , namely, a hypergeometric distribution $\mathbf{HG}(|N_s|, |N_k \cap N_s|, K)$. Recall that under the good event \mathcal{G}_s^2 , we have $(1 - \eta)sp \leq |N_s| \leq (1 + \eta)sp$ and $|N_k \cap N_s| \leq (\log n)^3$ for each $\lfloor \eta n \rfloor \leq k \leq s - 1$.

Now for any triple (N, M, K) with $(1 - \eta)sp \leq N \leq (1 + \eta)sp$, $0 \leq M \leq (\log n)^3$ and $F_s \leq K \leq 6F_s \ll \log n$, we have

$$\begin{aligned} \mathbb{P}[\mathbf{HG}(N, M, K) \geq 1] &= \sum_{k \geq 1} \frac{\binom{M}{k} \binom{N-M}{K-k}}{\binom{N}{K}} \leq \sum_{k \geq 1} \binom{M}{k} \cdot \frac{N^{K-k}}{(K-k)!} \cdot \frac{K!}{(N-K)^K} \\ &\leq \sum_{k \geq 1} M^k \left(1 - \frac{K}{N}\right)^{-K+k} \cdot \frac{K^k}{k!(N-K)^k} \\ &\leq \exp\left(\frac{K^2}{N}\right) \sum_{k \geq 1} \frac{1}{k!} \left(\frac{KM}{N-K}\right)^k \leq \frac{eKM}{N-K} \leq \frac{(\log n)^4}{np}. \end{aligned}$$

As a result,

$$\widehat{\mathbb{E}}_s \left[\sum_{k \geq 1} |R_s^k| \mid F_s \leq E_{r,s} \leq 6F_s \right] \leq (\log n)^4/p \ll n^{0.6}, \quad (5.26)$$

where the last inequality follows since we are under the assumption $p = n^{-1/2+o(1)}$. Similarly we have

$$\begin{aligned} \mathbb{P}[\mathbf{HG}(N, M, K) \geq 3] &= \sum_{k \geq 3} \frac{\binom{M}{k} \binom{N-M}{K-k}}{\binom{N}{K}} \leq \sum_{k \geq 3} \binom{M}{k} \left(1 - \frac{K}{N}\right)^{-K+k} \left(\frac{K}{N-K}\right)^k \\ &\leq \left(1 - \frac{K}{N}\right)^{-K} \sum_{k \geq 3} \frac{1}{k!} \left(\frac{MK}{N-K}\right)^k \leq \frac{e(KM)^3}{(N-K)^3} \leq \frac{(\log n)^{12}}{(np)^3}, \end{aligned}$$

thus

$$\widehat{\mathbb{E}}_s \left[\sum_{k \geq 3} |R_s^k| \mid F_s \leq E_{r,s} \leq 6F_s \right] = o(1). \quad (5.27)$$

Therefore, we see from Markov inequality that $|R_s^1| + |R_s^2| \leq n^{0.6}$ and $R_s^k = \emptyset$ for any $k \geq 3$ with high probability, finishing the proof of the lemma. \square

Recall that

$$\mathcal{A}_{r,s} = \bigcap_{k \in A_{r,s}} \{E_{r,k} < F_k\} = \bigcap_{k \in A_{r,s}} \left\{ \sum_{j \in N_k} G_{r,\pi^*}(j) < F_k \right\}.$$

By introducing these random sets R_s^k , we have

$$\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s} \mid F_s \leq E_{r,s} \leq 6F_s] = \widehat{\mathbb{P}}_s \left[\sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} < F_k - \ell, \forall \ell \geq 0, k \in R_s^\ell \mid F_s \leq E_{r,s} \leq 6F_s \right].$$

From the total probability formula, we see $\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s} \mid F_s \leq E_{r,s} \leq 6F_s]$ can be expressed as

$$\sum_{(R_s^0, R_s^1, \dots)} \widehat{\mathbb{P}}_s[\mathcal{R}_s = (R_s^0, R_s^1, \dots) \mid F_s \leq E_{r,s} \leq 6F_s] \times \\ \widehat{\mathbb{P}}_s \left[\sum_{j \in N_r \setminus N_s} G_{r,\pi^*(j)} < F_r - \ell, \forall \ell \geq 0, r \in R_s^\ell \mid F_s \leq E_{r,s} \leq 6F_s, \mathcal{R}_s = (R_s^0, R_s^1, \dots) \right],$$

where the summation is taken over all possible realizations of $\mathcal{R} = (R_s^0, R_s^1, \dots)$. Note that the event $\{F_s \leq E_{r,s} \leq 6F_s\}$ and the random sets $R_s^\ell, \ell \in \mathbb{N}$ are all measurable with respect to the random variables $\{G_{r,\pi^*(j)} : j \in N_s\}$, which are independent with the event in the second probability term, so the conditioning can be removed. Combining with Lemma 5.18, it suffices to show for any realization of $\mathcal{R}_s = (R_s^0, R_s^1, \dots)$ that is typical, say (R_s^0, R_s^1, \dots) , it always holds

$$\widehat{\mathbb{P}}_s \left[\sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} < F_k - \ell, \forall \ell \geq 0, k \in R_s^\ell \right] \geq [1 - o(1)] \widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}].$$

To tackle this relation, a trivial lower bound simplifies our goal to show

$$\widehat{\mathbb{P}}_s \left[\sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} < F_k - \ell, \forall \ell \geq 0, k \in R_s^\ell \mid \sum_{j \in N_k} G_{k,\pi^*(j)} < F_k \right] \geq 1 - o(1).$$

or equivalently (recall that $R_s^\ell = \emptyset$ for $\ell \geq 3$ since (R_s^0, R_s^1, \dots) is typical),

$$\widehat{\mathbb{P}}_s \left[\exists \ell \in \{1, 2\}, k \in R_s^\ell \text{ s.t. } \sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} \geq F_k - \ell \mid \mathcal{A}_{r,s} \right] \leq o(1). \quad (5.28)$$

Since the event in the probability is increasing and $\mathcal{A}_{r,s}$ is decreasing, from FKG inequality and a simple union bound we see the left hand side of (5.28) is bounded by

$$\begin{aligned} & \widehat{\mathbb{P}}_s \left[\exists \ell \in \{1, 2\}, k \in R_s^\ell \text{ s.t. } \sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} \geq F_k - \ell \right] \\ & \leq (|R_s^1| + |R_s^2|) \max_{k \in A_{r,s}} \widehat{\mathbb{P}}_s \left[\sum_{j \in N_k \setminus N_s} G_{k,\pi^*(j)} \geq F_k - 2 \right] \\ & \leq n^{0.6} \times n^{-0.99} = o(1), \end{aligned}$$

as desired. Where in the last inequality we used the assumption that (R_s^0, R_s^1, \dots) is typical and standard estimations of Binomial variables (see e.g. Lemma 5.24, note that $0 < \eta < 0.01$). This proves (5.28) and thus yields (5.24), which completes the proof of (5.22). Therefore, Proposition 5.13 has been established and the analysis for p in the sparse regime has been done.

5.3.3 Analysis for dense regime

In this subsection we give the analysis in the dense regime, i.e. $p_c \ll p \ll 1/(\log n)^3$. We will deal with the cases $np^3 \leq n^{-0.1}$ and $np^3 \geq n^{-0.1}$ separately, where the former case can be tackled almost as same as before (the previous approach essentially works whenever $np^3 \ll 1$), while for the latter case we will take another approach.

The case $np^3 \leq n^{-0.1}$ For parameter p in this regime, the proof essentially goes in a same way as the previous subsection, we just need to make corresponding verbal changes. Now let F_s be defined as in (5.19) where we pick $\varepsilon_n = \eta$. We define the good events \mathcal{G}_{s-1}^1 and \mathcal{G}_s^2 similarly as Definition 5.15.

Definition 5.3 (Good events in the dense regime). *Denote \mathcal{G}_{s-1}^1 for the event that*

$$(1 - \eta)rp \leq N_r \leq (1 + \eta)rp.$$

In addition, we let \mathcal{G}_s^2 be the event that $(1 - \eta)sp \leq N_s \leq (1 + \eta)sp$ and

$$|N_r \cap N_s| \leq (1 + \eta)rp^2, \forall [\eta n] \leq r \leq s - 1. \quad (5.29)$$

Similar as the proof of Proposition 5.16 we can show that $\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1)$ and $\mathbb{P}[\mathcal{G}_s^2 | \mathcal{F}_{s-1}] = 1 - o(1)$ whenever \mathcal{F}_{s-1} satisfies \mathcal{G}_{s-1}^1 . We will prove an analogue of Proposition 5.17 for p in this regime, namely for any \mathcal{F}_{s-1} satisfies \mathcal{G}_{s-1}^1 and \mathcal{N}_s satisfies \mathcal{G}_s^2 , it holds that

$$\mathbb{E}[X_s | \mathcal{F}_{s-1}, \mathcal{N}_s] \geq (1 - o(1)) \widehat{\mathbb{E}}_s[X_s] \rightarrow \infty,$$

and

$$\mathbb{E}[X_s^2 | \mathcal{F}_{s-1}, \mathcal{N}_s] \leq (1 + o(1)) \left(\widehat{\mathbb{E}}_s[X_s] \right)^2.$$

Again it is straightforward to check that $\widehat{\mathbb{E}}_s[X_s] \rightarrow \infty$ (see e.g. Lemma 5.25) and thus the second moment estimate follows in a same way as before. In order to obtain the first moment estimate, recall the definition of the random set of indices $R_s^\ell, \ell \in \mathbb{N}$, and now we call a realization (R_s^0, R_s^1, \dots) of \mathcal{R}_s to be typical, if

$$\sum_{\ell=1}^{10} |R_s^\ell| \leq n^{0.95},$$

and $R_s^\ell = \emptyset, \forall \ell \geq 11$. Following the arguments given as before, we see it suffices to show for any fixed $r \in R_{s-1}$, the following three things hold:

$$(i) \quad \widehat{\mathbb{P}}_s[F_s \leq E_{r,s} \leq 2sp^2] = [1 - o(1)]\widehat{\mathbb{P}}_s[E_{r,s} \geq F_s].$$

$$(ii) \quad \mathbb{P}[\mathcal{R}_s \text{ is typical} \mid F_s \leq E_{r,s} \leq 2sp] = 1 - o(1).$$

(iii) Under any typical realization (R_s^0, R_s^1, \dots) of \mathcal{R}_s ,

$$\widehat{\mathbb{P}}_s \left[\sum_{j \in N_k \setminus N_s} \mathbf{G}_{k, \pi^*(j)} < F_k - \ell, \forall \ell \geq 0, k \in R_s^\ell \mid \sum_{j \in N_k} \mathbf{G}_{k, \pi^*(j)} < F_k, \forall k \in A_{r,s} \right] \geq 1 - o(1).$$

(i) follows readily from the tail estimations of binomial variables. For (ii), we recall that conditioned on the event $\{E_{r,s} = K\}$, then for each $k \in A_{r,s}$, $\sum_{j \in N_r \setminus N_s} \mathbf{G}_{k, \pi^*(j)}$ has distribution $\mathbf{HG}(|N_s|, |N_s \cap N_k|, K)$. Under the good events, we have $(1-\eta)sp \leq |N_s| \leq (1+\eta)sp$, $|N_s \cap N_k| \leq (1+\eta)rp^2$. Note that for any triple (N, M, K) with $(1-\eta)sp \leq N \leq (1+\eta)sp$, $M \leq (1+\eta)sp^2$ and $F_s \leq K \leq 2sp^2$, we have

$$\mathbb{P}[\mathbf{HG}(N, M, K) \geq 1] \leq \sum_{k \geq 1} \frac{1}{k!} \left(\frac{MK}{N-K} \right)^k \leq \frac{eMK}{N-K} \leq \frac{2e(1+\eta)(sp^2)^2}{(1-\eta)sp - 2sp^2} \ll n^{-0.05},$$

and

$$\mathbb{P}[\mathbf{HG}(N, M, K) \geq 11] \leq \sum_{k \geq 11} \frac{1}{k!} \left(\frac{MK}{N-K} \right)^k \leq e \left(\frac{MK}{N-K} \right)^{11} \ll n^{-1},$$

so (ii) follows from Markov inequality. Finally for (iii), it is equivalent to

$$\widehat{\mathbb{P}}_s \left[\exists \ell \in \{1, \dots, 10\}, k \in R_s^\ell, \sum_{j \in N_k \setminus N_s} \geq F_k - \ell \mid \mathcal{A}_{r,s} \right] = o(1),$$

which can be again shown by a combination with the union bound and tail estimations for Binomial variables. This completes the analysis for $p \gg p_c$ and $np^3 \leq n^{-0.1}$.

The case $np^3 \geq n^{-0.1}$ Now we consider the case when p satisfies $np^3 \geq n^{-0.1}$ and $p \ll 1/(\log n)^3$. We let F_s defined as in (5.19) where we take $\varepsilon_n = \eta$ if $p \leq n^{-0.1}$ and $\varepsilon_n = \log \log n / \sqrt{\log n}$ if $n^{-0.1} \leq p \leq 1/(\log n)^3$. For p in such a regime, the previous approach will not always work, so we turn to another approach which exploits the pure randomness of N_s given \mathcal{F}_{s-1} . For each $r \in R_{s-1}$, we write $N_s(r) = \{j < s : \mathbf{G}_{r, \pi^*(j)} = 1\}$ for the set of indices j such that $\mathbf{v}_{\pi^*(j)}$ is a neighbor of \mathbf{v}_r in the matched vertices. Recall the definition of \tilde{N}_s and note that when conditioned on \mathcal{F}_{s-1} , the random sets $N_s(r), r \in R_{s-1}$ are all measurable with respect to \tilde{N}_s . We also note that conditioned on any realization of $\mathcal{F}_{s-1} \cup \tilde{N}_s$,

$$E_{r,s} = \sum_{j < s} G_{s,j} \mathbf{G}_{r, \pi^*(j)} = \sum_{j \in N_s(r)} G_{s,j},$$

which has binomial distribution $\mathbf{B}(|\mathbf{N}_s(r)|, p)$ since the random variables $\{G_{s,j} : j < s\}$ are independent with \mathcal{F}_{s-1} and $\tilde{\mathcal{N}}_s$.

As before, we start by presenting the definitions of good events for p in this regime.

Definition 5.4. Denote \mathcal{G}_{s-1}^1 for the event that

$$|\mathbf{N}_i| \leq ip + \sqrt{8ip(1-p)\log n}, \forall \lfloor \eta n \rfloor + 1 \leq i \leq s-1.$$

In addition, we write $\delta_n = 0.1$ for $p \leq n^{-0.1}$ and $\delta_n = \frac{\log \log n + \log \log \log n}{\log n}$ for $n^{-0.1} \leq p \leq 1/(\log n)^3$, then denote \mathcal{G}_s^2 for the event that the following two items hold:

- $sp - \sqrt{sp(1-p)n^{\delta_n}} \leq |\mathbf{N}_s(r)| \leq sp + \sqrt{sp(1-p)n^{\delta_n}}$ for all $r \in \mathbf{R}_{s-1}$.
- $|\mathbf{N}_s(i) \cap \mathbf{N}_s(j)| \leq 2sp^2$ for all $i, j \in \mathbf{R}_{s-1}$.

Note that given \mathcal{F}_{s-1} , \mathcal{G}_s^2 is measurable with respect to $\tilde{\mathcal{N}}_s$. The next lemma clarifies that both good events are typical.

Lemma 5.19. It holds that $\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1)$ and for any realization of \mathcal{F}_{s-1} that satisfies \mathcal{G}_{s-1}^1 , one has $\mathbb{P}[\mathcal{G}_s^2 \mid \mathcal{F}_{s-1}] = 1 - o(1)$.

Here we just sketch the proof of this lemma and postpone the precise argument to the appendix. $\mathbb{P}[(\mathcal{G}_{s-1}^1)^c] = o(1)$ is easy, and to show $\mathbb{P}[(\mathcal{G}_s^2)^c \mid \mathcal{F}_{s-1}] = o(1)$, the main difficulty lies in dealing with the conditioning. It is clear that $(\mathcal{G}_s^2)^c$ is measurable with respect to \mathbf{E}_s so conditioning on \mathcal{F}_{s-1} is equivalent to conditioning on $\mathcal{D}_{s-1} = \bigcap_{r' \in \mathbf{R}_{s-1}} \mathcal{A}_{r',s}$, which is a decreasing event. As a result, for the increasing part of $(\mathcal{G}_s^2)^c$, we can apply FKG inequality to remove the conditioning and then take a union bound. As for the decreasing part of $(\mathcal{G}_s^2)^c$, i.e. $\mathcal{D}_s^2 = \{|\mathbf{N}_s(r)| \leq sp - \sqrt{sp(1-p)n^{\delta_n}}\}$, we note that it is independent with $\{\mathcal{A}_{r',s}, r' \in \mathbf{R}_{s-1}, r' \neq r\}$, thus we have

$$\mathbb{P}[\mathcal{D}_s^2 \mid \mathcal{F}_{s-1}] = \widehat{\mathbb{P}}_s[\mathcal{D}_s^2 \mid \mathcal{D}_{s-1}] = \widehat{\mathbb{P}}_s[\mathcal{D}_s^2 \mid \mathcal{A}_{r,s}] \leq \widehat{\mathbb{P}}_s[\mathcal{D}_s^2] / \widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}], \quad (5.30)$$

and our choices of ε_n and δ_n will ensure the last quantity is $o(1)$.

Given these results, it remains to show the following conditional moment estimations.

Proposition 5.20. For any realization of \mathcal{F}_{s-1} such that \mathcal{G}_{s-1}^1 holds and any realization of $\tilde{\mathcal{N}}_s$ such that \mathcal{G}_s^2 holds, one has

$$\mathbb{E}[X_s^2 \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s] \leq [1 + o(1)] (\mathbb{E}[X_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s])^2. \quad (5.31)$$

The complete proof of Proposition 5.20 is also deferred into the appendix, here we just give some remarks. Since the random variables $\{G_{s,j} : j < s\}$ are independent of $\sigma(\mathcal{F}_{s-1} \cup \tilde{\mathcal{N}}_s)$, the computation of the conditional first and second moments are simply computations for (correlated) Binomial variables. The good event \mathcal{G}_s^2 controls the correlation structure and guarantees the correlation is weak enough so that (5.31) holds. It is also worthy to point out that in some detailed computations, we will need the assumption $np^3 \geq n^{-0.1}$ and this is why we need to deal with the case $np^3 \leq n^{-0.1}$ separately as before. This finishes the analysis of p satisfying $np^3 \geq n^{-0.1}$ and $p \ll 1/(\log n)^3$.

5.4 Computational hardness for online algorithms

In this section we prove the hardness result for online algorithms, i.e. the second item in Theorem 5.4. We start by presenting the definition of the set \mathcal{H} as below.

Definition 5.5. *Let \mathcal{H} be the set of simple graphs G on n vertices which satisfies*

$$\left| E(G) - \binom{n}{2} p \right| \leq \frac{\varepsilon}{3} D_{n,p}, \quad (5.32)$$

and for each induced subgraph H of G with $k \geq \alpha_1 n$ vertices, the number of edges in H satisfies

$$\left| E(H) - \binom{k}{2} p \right| \leq \frac{n^2 p}{\log n^{1/4}}. \quad (5.33)$$

We show that an Erdős-Rényi graph $G \sim \mathbf{G}(n, p)$ belongs to \mathcal{H} with high probability.

Lemma 5.21. *For $G \sim \mathbf{G}(n, p)$, $\mathbb{P}[G \in \mathcal{H}] = 1 - o(1)$.*

Proof. It is clear that $|E(G)| \sim \mathbf{B}(\binom{n}{2}, p)$ and thus by the Chernoff bounds (5.6) and (5.5), we see

$$\mathbb{P} \left[\left| E(G) - \binom{n}{2} p \right| \geq \frac{\varepsilon}{3} D_{n,p} \right] \leq 2 \exp \left(-\frac{\varepsilon^2 n^3 p^2 \log n}{9(n^2 p + \varepsilon D_{n,p})} \right) = o(1),$$

thus (5.32) holds with high probability. To show (5.33) is true for every H with $|V(H)| \geq \alpha_1 n$ with high probability, we note that for each such induced subgraph, the distribution of number of edges in H is given by $\mathbf{B}(\binom{k}{2}, p)$. By applying the Chernoff bound we get for n large enough,

$$\mathbb{P} \left[\left| E(H) - \binom{k}{2} p \right| \geq \frac{n^2 p}{\log n^{1/4}} \right] \leq 2 \exp \left(-\frac{n^4 p^2}{2n^2 p \log n^{1/2}} \right) \leq \exp(-n^{3/2}),$$

where the last inequality is because $p \gg p_c = \sqrt{\log n/n}$. The result then follows by a union bound since there are at most $2^n = \exp(O(n))$ subgraphs H in total. \square

In the remaining of this section, we will fix some arbitrary graph $G \in \mathcal{H}$.

5.4.1 The forbidden structure

Fix $\varepsilon > 0$. The goal of this subsection is to design certain forbidden structure and show it is very unlikely to exist. Note that by the definition of Riemann integral, we have for a sequence $0 = \alpha_0 < \alpha_1 < \dots < \alpha_N = 1$ with $\Delta = \max_{1 \leq i \leq N} (\alpha_i - \alpha_{i-1})$,

$$\lim_{\Delta \rightarrow 0} \sum_{i=1}^N (\alpha_i - \alpha_{i-1}) \sqrt{\alpha_i + \alpha_{i-1}} = \int_0^1 \sqrt{2x} \, dx = \frac{2\sqrt{2x^3}}{3} \Big|_0^1 = \beta_c.$$

Therefore, we can pick some integer N together with $0 = \alpha_0 < \alpha_1 < \dots < \alpha_N = 1$ such that

$$\sum_{i=1}^N (\alpha_i - \alpha_{i-1}) \sqrt{\alpha_i + \alpha_{i-1}} < \beta_c + \varepsilon/3. \quad (5.34)$$

Denote $\frac{\beta_c + 2\varepsilon/3}{\beta_c + \varepsilon/3}$ by $1 + \delta$. Fix a large integer D with

$$D > \max_{1 \leq k \leq N} \frac{\alpha_{k-1}}{2\delta(\alpha_k - \alpha_{k-1})}. \quad (5.35)$$

Consider a D regular tree \mathcal{T} with $(N+1)$ generations, and denote the set of leaves in \mathcal{T} by \mathcal{L} , write $L = |\mathcal{L}| = D^N$. For a vertex v of \mathcal{T} , let $|v|$ denote for its depth in \mathcal{T} . If $v \in \mathcal{L}$, let $v(k)$ be the ancestor of v with depth k and for another $w \in \mathcal{L}$, we let $v \wedge w$ be the common ancestor of v and w with largest depth. We define a set of correlated instances as follow:

Definition 5.6 (Correlated instances). *For each vertex v of \mathcal{T} , assign it with independent Bernoulli variables $\{E_{i,j}^v\}_{1 \leq i < j \leq \alpha_{|v|}n}$. A correlated instances is a set of graphs $\{G^v\}_{v \in \mathcal{L}}$ indexed by \mathcal{L} , where for each $v \in \mathcal{L}$, the adjacency matrix of G^v is defined by*

$$G_{i,j}^v = E_{i,j}^{v(k)}, \text{ if } \alpha_{k-1}n < j \leq \alpha_k n, \forall i < j.$$

Now we are able to state the forbidden structure for correlated instances.

Definition 5.7 (Forbidden structure for correlated instances). *For a correlated instances $\{G^v\}_{v \in \mathcal{L}}$, we say a set of permutations $\{\pi_v\}_{v \in \mathcal{L}}$ has the forbidden structure, if*

- (i) *For each $v \in \mathcal{L}$, $\pi_v \in S_{\beta_c + \varepsilon}(G^v, \mathbf{G})$.*
- (ii) *For each pair of $u, v \in \mathcal{L}$, it holds that $\pi_u(i) = \pi_v(i)$ for any $1 \leq i \leq \alpha_{|u \wedge v|}n$.*

We will show in the next proposition that for any $\mathbf{G} \in \mathcal{H}$, the forbidden structure exists with probability no more than $\exp(-\Omega(n \log n))$. This will be the crucial ingredient for deriving the hardness results for online algorithms.

Proposition 5.22. *There exists $c > 0$ such that for any $\mathbf{G} \in \mathcal{H}$,*

$$\mathbb{P}[\text{the forbidden structure exists}] \leq \exp(-cn \log n), \quad (5.36)$$

where the probability is taken over correlated instances generated as in Definition 5.6.

Proof. We condition on the realization of \mathbf{G} and assume that \mathcal{H} holds. Assume there exists such forbidden structure $\{\pi_v\}_{v \in \mathcal{L}}$. We define a set of embeddings $\{\sigma_v\}_{v \in \mathcal{T}}$ indexed by the vertices of \mathcal{T} as follows: for each vertex v of \mathcal{T} , σ_v is a map from $\{i : \alpha_{|v|-1}n < i \leq \alpha_{|v|}n\}$ to $[n]$ with $\sigma_v(i) = \pi_{v'}(i)$ for each i , where v' is any descendant of v in \mathcal{L} . Note that condition (ii) guarantees $\{\sigma_v\}_{v \in \mathcal{L}}$ is well-defined, and the mapping $\{\pi_v\}_{v \in \mathcal{L}} \mapsto \{\sigma_v\}_{v \in \mathcal{T}}$ is clearly injective.

For each $v \in \mathcal{T}$, let γ_v such that

$$\sum_{\substack{i < j, \\ 1 \leq i \leq \alpha_{|v|}n, \\ \alpha_{|v|-1}n < j \leq \alpha_{|v|}n}} G_{i,j}^v \mathbf{G}_{\sigma_v(i), \sigma_v(j)} = \sum_{\substack{i < j, \\ 1 \leq i \leq \alpha_{|v|}n, \\ \alpha_{|v|-1}n < j \leq \alpha_{|v|}n}} \mathbf{G}_{\sigma_v(i), \sigma_v(j)} p + \gamma_v D_{n,p}. \quad (5.37)$$

Then condition (i) translates to the following:

$$\sum_{k=0}^N \gamma_{v(k)} \geq \beta_c + \varepsilon + \frac{|\binom{n}{2}p - |E(\mathbf{G})||}{D_{n,p}} \stackrel{(5.32)}{\geq} \beta_c + 2\varepsilon/3, \quad \forall v \in \mathcal{L}.$$

Summing over the above inequality for all $v \in \mathcal{L}$ we get

$$\sum_{k=0}^N \frac{1}{D^k} \sum_{|v|=k} \gamma_v \geq \beta_c + 2\varepsilon/3.$$

Combining with (5.34), we see there must exists some $0 \leq k \leq N$ such that

$$\frac{1}{D^k} \sum_{|v|=k} \gamma_v \geq \frac{\beta_c + 2\varepsilon/3}{\beta_c + \varepsilon/3} \cdot (\alpha_k - \alpha_{k-1}) \sqrt{\alpha_k + \alpha_{k-1}} = (1 + \delta)(\alpha_k - \alpha_{k-1}) \sqrt{\alpha_k + \alpha_{k-1}}. \quad (5.38)$$

and we denote this event as $\mathcal{E}_k, k = 0, 1, \dots, N$. We conclude that

$$\mathbb{P}[\text{forbidden structure exists}] \leq \sum_{k=0}^N \mathbb{P}[\mathcal{E}_k].$$

Now it remains to bound the probability of \mathcal{E}_k and we shall do this by the first moment method. Note that \mathcal{E}_k only depends on those σ_v with $|v| \leq k$, so we have $\mathbb{P}[\mathcal{E}_k]$ is bounded by the number of possible realizations of $\{\sigma_v\}_{|v| \leq k}$ times the probability that (5.38) happens for some fixed $\{\sigma_v\}_{|v| \leq k}$. It is not hard to see that the possible realization of such $\{\sigma_v\}_{|v| \leq k}$ is bounded by

$$\exp \left(\sum_{i=1}^k D^{i-1} (\alpha_i - \alpha_{i-1}) n \log n + O(n) \right).$$

In order to control the probability term, we note that by our constructions, the family of random variables

$$X_v^\sigma \stackrel{\text{def}}{=} \sum_{\substack{i < j, \\ 1 \leq i \leq \alpha_{|v|} n, \\ \alpha_{|v|-1} n < j \leq \alpha_{|v|} n}} G_{i,j}^v \mathbf{G}_{\sigma_v(i), \sigma_v(j)}, v \in \mathcal{T}$$

are mutually independent, while each X_v^σ has the distribution of $\mathbf{B}(N_v^\sigma, p)$, where

$$N_v^\sigma \stackrel{\text{def}}{=} \sum_{\substack{i < j, \\ 1 \leq i \leq \alpha_{|v|} n, \\ \alpha_{|v|-1} n < j \leq \alpha_{|v|} n}} \mathbf{G}_{\sigma_v(i), \sigma_v(j)}.$$

Under the event \mathcal{H} we have from (5.33) that

$$N_v^\sigma = \frac{\alpha_{|v|}^2 - \alpha_{|v|-1}^2}{2} n^2 p + o(n^2 p),$$

so by the Chernoff bound we see for each vertex $v \in \mathcal{T}$ and any constant γ_v ,

$$\begin{aligned} \mathbb{P}[X_v^\sigma \geq N_v^\sigma p + \gamma_v D_{n,p}] &\leq \exp \left(- \frac{(\gamma_v \vee 0)^2 n^3 p^2 \log n}{2(N_v^\sigma p + (\gamma_v \vee 0) D_{n,p})} \right) \\ &\leq \exp \left(- \frac{(\gamma_v \vee 0)^2}{\alpha_{|v|}^2 - \alpha_{|v|-1}^2} n \log n + o(n \log n) \right). \end{aligned}$$

Write $M = \sqrt{D^{k-1}(1+\delta)^2(\alpha_k - \alpha_{k-1})}$, and we denote

$$\Gamma_k = \left\{ (\gamma_v)_{|v| \leq k} : \sum_{|v|=k} \gamma_v \geq (1+\delta)(\alpha_k - \alpha_{k-1}) \sqrt{\alpha_k + \alpha_{k-1}}, \gamma_v \leq M, \forall |v| \leq k \right\}.$$

First by the above estimation we see that the probability that $X_v^\sigma \geq N_v^\sigma + M D_{n,p}$ for some $v \in \mathcal{T}$ is no more than $\exp(-M^2 n \log n + o(n \log n))$. As a result, we see the probability term is bounded by $\exp(-M^2 n \log n + o(n \log n))$ plus

$$\exp(O(\log n)) \times \sup_{(\gamma_v)_{|v| \leq k} \in \Gamma_k} \mathbb{P}[X_v^\sigma \geq N_v^\sigma + \gamma_v D_{n,p}, \forall v \text{ s.t. } |v| \leq k].$$

As a result of independence, the supremum of the probability term above is bounded by

$$\begin{aligned}
& \sup_{(\gamma_v)_{|v| \leq k} \in \Gamma_k} \exp \left(- \sum_{i=1}^k \sum_{|v|=i} \frac{(\gamma_v \vee 0)^2}{\alpha_i^2 - \alpha_{i-1}^2} n \log n + o(n \log n) \right) \\
& \leq \sup_{(\gamma_v)_{|v| \leq k} \in \Gamma_k} \exp \left(- \sum_{i=1}^k D^{i-1} \frac{(\sum_{|v|=i} \gamma_v \vee 0)^2}{\alpha_i^2 - \alpha_{i-1}^2} n \log n + o(n \log n) \right) \\
& \leq \sup_{(\gamma_v)_{|v| \leq k} \in \Gamma_k} \exp \left(- D^{k-1} \frac{(\sum_{|v|=k} \gamma_v \vee 0)^2}{\alpha_k^2 - \alpha_{k-1}^2} n \log n + o(n \log n) \right) \\
& \leq \exp \left(- D^{k-1} (1 + \delta)^2 (\alpha_k - \alpha_{k-1}) n \log n + o(n \log n) \right),
\end{aligned}$$

where the first inequality comes from Cauchy-Schwartz and the last inequality follows from the definition of Γ_k . With these estimates in hand, we get $\mathbb{P}[\mathcal{E}_k]$ is bounded by

$$\begin{aligned}
& \exp \left(\sum_{i=1}^k D^{i-1} (\alpha_i - \alpha_{i-1}) n \log n + o(n \log n) \right) \times \\
& \quad \left[\exp \left(- M^2 n \log n + o(n \log n) \right) + \exp \left(- D^{k-1} (1 + \delta)^2 (\alpha_k - \alpha_{k-1}) n \log n + o(n \log n) \right) \right] \quad (5.39) \\
& \leq 2 \exp \left([D^{k-2} \alpha_{k-1} - 2\delta D^{k-1} (\alpha_k - \alpha_{k-1})] n \log n + o(n \log n) \right) = o(1),
\end{aligned}$$

as desired (the last inequality is true from our choice of D in (5.35)). \square

5.4.2 Completing the proof

Let \mathcal{A} be an arbitrary online algorithm. By definition, we can assume that there exist probability spaces $\Omega_1, \dots, \Omega_n$ and deterministic functions $f_k : \Omega_k \rightarrow [n]$, such that \mathcal{A} operates as follows: for each $1 \leq k \leq n$, assuming that $\pi^*(1), \dots, \pi^*(k-1)$ are determined, \mathcal{A} samples $\omega_k \in \Omega_k$ according to some probability measure \mathbb{P}_k determined by $\{G_{ij}\}_{1 \leq i < j \leq k}$, $\{\mathbf{G}_{ij}\}_{1 \leq i < j \leq n}$ and $\pi^*(1), \dots, \pi^*(k-1)$, and sets $\pi^*(k) = f_k(\omega_k)$. The mechanism should also ensure that the final output π^* is almost surely a permutation in S_n .

Still we fix some $\mathbf{G} \in \mathcal{H}$. For correlated instances of graphs $\{G^v\}_{v \in \mathcal{L}}$ as defined in Definition 5.6, we run \mathcal{A} simultaneously on the pairs $(G^v, \mathbf{G}), v \in \mathcal{L}$ with outputs $\{\pi_v^*\}_{v \in \mathcal{L}}$, where the sampling procedure obeys the following rule: for any two vertices $u, v \in \mathcal{L}$, the internal randomness ω_k for $\pi_u^*(k)$ and $\pi_v^*(k)$ are identically sampled if $k \leq \alpha_{|u \wedge v|} n$, while they are independent sampled if $k > \alpha_{|u \wedge v|} n$. Note that by our construction of correlated instances and the definition of online algorithms, one can prove by induction that the law \mathbb{P}_k of $\pi_u^*(k)$ and $\pi_v^*(k)$ are the same as long as

$1 \leq k \leq \alpha_{|u \wedge v|} n$. This demonstrates that the sampling rule above is self-consistent and that such a running procedure is valid.

We have the following lower bound for the probability of $\pi_v^* \in S_{\beta_c + \varepsilon}(G, \mathbb{G})$ for all $v \in \mathcal{L}$.

Proposition 5.23. *Assume that the output π^* of \mathcal{A} satisfies*

$$\mathbb{P}[\pi^* \in S_{\beta_c + \varepsilon}(G, \mathbb{G})] = \mathbb{E}_{G \sim \mathbf{G}(n, p)} \mathbb{Q}[\pi^* \in S_{\beta_c + \varepsilon}(G, \mathbb{G})] \geq p_{\text{suc}},$$

where \mathbb{Q} is the internal randomness of \mathcal{A} . Then through the aforementioned procedure, we have

$$\mathbb{P}[\pi_v^* \in S_{\beta_c + \varepsilon}(G^v, \mathbb{G}), \forall v \in \mathcal{L}] \geq p_{\text{suc}}^L, \quad (5.40)$$

Proof. Denote S_v as the event that $\pi_v^* \in S_{\beta_c + \varepsilon}(G^v, \mathbb{G})$. First we conditioned on \mathbb{G} . For $i = 0, 1, \dots, N$, define \mathcal{F}_i as the σ -field generated by $\{G_{ij}^v\}_{1 \leq i < j \leq \alpha_i n, \forall v \in \mathcal{L} \text{ and } \pi^*(k), 1 \leq k \leq \alpha_i n}$. For each $v \in \mathcal{T}$, we use $\mathcal{L}(v)$ to denote the offspring of v in \mathcal{L} and we write $\mathcal{S}(v)$ for the event that $\pi_u^* \in S_{\beta_c + \varepsilon}(G, \mathbb{G})$ for each $u \in \mathcal{L}(v)$. Note that we are interested in the event that $\mathbb{P}[\mathcal{S}(v_o)]$ for the root v_o , and we will show by induction that

$$\mathbb{P}[\mathcal{S}(v_o)] \geq \mathbb{E} \prod_{|v|=i} \mathbb{P}[\mathcal{S}(v)], \forall 0 \leq i \leq N. \quad (5.41)$$

The case $i = 0$ is trivial. Assume (5.41) holds for some $0 \leq i \leq N - 1$, we note that for each v with $|v| = i$, denote its D decedents by v_1, \dots, v_D , then $\mathcal{S}(v) = \bigcap_{k=1}^D \mathcal{S}(v_k)$ and the events $\mathcal{S}(v_k), k = 1, 2, \dots, D$ are conditional independent given \mathcal{F}_i . By conditioning on \mathcal{F}_{i+1} , we see from the iterated expectation that

$$\begin{aligned} \mathbb{P}[\mathcal{S}(v)] &= \mathbb{E} \left[\prod_{k=1}^D \mathbb{P}[\mathcal{S}(v_k) \mid \mathcal{F}_i] \right] = \mathbb{E} \left[\mathbb{P}[\mathcal{S}(v_1) \mid \mathcal{F}_i]^D \right] \\ &\geq \left(\mathbb{E} [\mathbb{P}[\mathcal{S}(v_1) \mid \mathcal{F}_i]] \right)^D = (\mathbb{P}[\mathcal{S}(v_1) \mid \mathbb{G}])^D = \prod_{k=1}^D \mathbb{P}[\mathcal{S}(v_k)]. \end{aligned}$$

where we used the symmetry between v_1, \dots, v_D twice and the inequality follows from Jensen's inequality. Thus by taking product of all v with $|v| = i$ we see (5.41) holds with $i + 1$. This establishes (5.41) for each $0 \leq i \leq N$ and in particular, we see for $i = N$,

$$\mathbb{P}[\mathcal{S}(v_o)] \geq \prod_{|v|=N} \mathbb{P}[\mathcal{S}(v)] = (\mathbb{P}[\pi^* \in S_{\beta_c + \varepsilon}(G, \mathbb{G})])^L \geq p_{\text{suc}}^L,$$

completing the proof. \square

Note that if $\pi_v^* \in S_{\beta_c+\varepsilon}(G^v, \mathbf{G})$ for each $v \in \mathcal{L}$, then $(\pi_v^*)_{v \in \mathcal{L}}$ formulates the forbidden structure. Combining with Proposition 5.22, we get whenever \mathbf{G} belongs to \mathcal{H} , then for any $\mathcal{A} \in \text{OGAA}$, it holds that

$$\mathbb{P}[\mathcal{A}(G, \mathbf{G}) \in S_{\beta_c+\varepsilon}(G, \mathbf{G})] \leq \exp(-cn \log n/L) = \exp(-c_0 n \log n),$$

thus by Markov property, we see

$$\mathbb{P}[G : \mathbb{Q}[\mathcal{A}(G, \mathbf{G}) \in S_{\beta_c+\varepsilon}(G, \mathbf{G}) \geq \exp(-c_0 n \log n/2)]] \leq \exp(-c_0 n \log n/2).$$

Note that this holds for any $\mathbf{G} \in \mathcal{H}$ and any $\mathcal{A} \in \text{OGAA}$, so the proof of the second item in Theorem 5.4 is completed.

5.5 Complimentary proofs

5.5.1 Tail Bounds

In this section, we present several tail estimates for binomial variables which will be necessary for our proofs. The tail bounds are divided into two parts according to different approximation regimes: the Poisson regime and the Normal regime. For ease of notation, we let \mathbf{B} , \mathbf{HG} and \mathbf{Poi} be binomial, hypergeometric and Poisson distribution and their corresponding random variables. In the subsequent lemmas, we will frequently use the following easy-checked estimation of binomial numbers:

$$\frac{n^k}{k!} \exp\left(-\frac{k^2}{n}\right) \leq \binom{n}{k} \leq \frac{n^k}{k!}. \quad (5.42)$$

Lemma 5.24 (Tail estimates for sparse regime). *For $\log n/n \ll p \ll p_c$ and an integer N with $np/2 \leq N \leq 2np$, let $X \sim \mathbf{B}(N, p)$. Then for any constant $\alpha > 0$,*

$$\mathbb{P}\left[X \geq \frac{\alpha \log n}{\log\left(\frac{\log n}{Np}\right)}\right] = n^{-\alpha+o(1)}.$$

Proof of lemma 5.24. For the upper bound, by applying the Chernoff bound given as in (5.5), we have the deviation probability is bounded by

$$\begin{aligned} & \exp\left(-Np \left(\left(\frac{\alpha \frac{\log n}{Np}}{\log\left(\frac{\log n}{Np}\right)}\right) \log\left(\frac{\alpha \frac{\log n}{Np}}{\log\left(\frac{\log n}{Np}\right)}\right) - \left(\left(\frac{\alpha \frac{\log n}{Np}}{\log\left(\frac{\log n}{Np}\right)}\right) - 1\right)\right)\right) \\ & \leq \exp\left(-Np \left(\alpha \frac{\log n}{Np} + o\left(\frac{\log n}{Np}\right)\right)\right) = n^{-\alpha+o(1)}. \end{aligned}$$

For the lower bound, it suffices to derive an one-point estimation. For simplicity we write $k = \lceil \frac{\alpha \log n}{\log(\frac{\log n}{Np})} \rceil$, and we have

$$\begin{aligned}
\mathbb{P}[X = k] &= \binom{N}{k} p^k (1-p)^{N-k} = \frac{(Np)^k}{k!} \prod_{i=1}^{k-1} \left(1 - \frac{i}{N}\right) (1-p)^{N-k} \\
&= (1 + o(1)) \frac{(Np)^k}{k!} \exp(-Np) \\
&= (1 + o(1)) \exp\left(k \log(Np) - k \log k - k + \log(\sqrt{2\pi k}) - Np\right) \\
&= (1 + o(1)) \exp\left(-k \log\left(\frac{k}{Np}\right) + o(\log n)\right) \\
&= (1 + o(1)) \exp(-\alpha \log n + o(\log n)).
\end{aligned}$$

The proof is completed. \square

The following well-known result gives general lower bounds for tail probabilities of Binomial variables. The complete proof can be found in e.g. [35, Theorem 2.1].

Lemma 5.25 (Lower bounds for binomial tail estimates). *If $0 \leq p \leq 1/4$, $np \leq k \leq n$ or $np \leq k \leq n(1-p)$, then for $X \sim \mathbf{B}(n, p)$,*

$$\mathbb{P}[X \geq k] \geq 1 - \Phi\left((k - np)/\sqrt{np(1-p)}\right).$$

where $\Phi(\cdot)$ is the distribution function of standard normal distribution $\mathcal{N}(0, 1)$.

Combining such lower bound with the Chernoff bound, we can derive the following estimates for the tail probabilities we are interested in:

Lemma 5.26 (Tail estimate for dense regime). *For an integer N and a parameter p , we have for any constant $\alpha > 0$, the following two tail bounds*

$$\mathbb{P}\left[|\mathbf{B}(N, p) - Np| > \sqrt{2\alpha Np(1-p) \log n}\right] \leq n^{-\alpha+o(1)},$$

hold under the assumption that either $p \gg p_c$, $N \geq \eta np/2$ or $np^3 \geq (\log n)^2$, $N \geq \eta np^2/2$.

Proof of lemma 5.26. Under the assumptions it holds $Np \gg \sqrt{Np \log n}$ and thus the result follows from (5.7). \square

Next we introduce an lemma on the upper bound on the lower tail probability,

Lemma 5.27. For binomial random variable $\mathbf{B}(n, p)$, we have,

- If $\sqrt{\log n/n} \ll p \leq n^{-0.1}$, take $\delta_n = 0.1$, then we have,

$$\mathbb{P} \left[\mathbf{B}(n, p) < np - \sqrt{np(1-p)n^{0.1}} \right] \leq \exp(-\Omega(n^{0.1})).$$

- If $n^{-0.1} \leq p \leq 1/(\log n)^3$, we take $\delta_n = \frac{\log \log n + \log \log \log n}{\log n}$,

$$\mathbb{P} \left[\mathbf{B}(n, p) < np - \sqrt{2np(1-p)n^{\delta_n}} \right] \leq \exp(-(1+o(1)) \log n \log \log n).$$

- When $n^{-0.1} \leq p \leq 1/(\log n)^3$, take $\varepsilon_n = \frac{\log \log n}{\sqrt{\log n}}$, for $N = O(np)$, we have,

$$\mathbb{P} \left[\mathbf{B}(N, p) \geq Np + \sqrt{2(1-2\varepsilon_n)Np(1-p) \log n} \right] \leq n^{-1+3\varepsilon_n}.$$

Proof of lemma 5.27. For the first two claim we use (5.7) in the same manner as the proof in lemma 5.26. For the third conclusion give a detailed calculation, which uses Theorem 1 in [2]. Let $\mathcal{H}(a|p) := a \log(1/p) + (1-a) \log((1-a)/(1-p))$. For simplicity we let $\beta = 1 - 2\varepsilon_n$,

$$\begin{aligned} \mathcal{H} \left(p + \sqrt{2\beta \frac{p(1-p) \log n}{N}} \middle| p \right) &= \left(p + \sqrt{2\beta \frac{p(1-p) \log n}{N}} \right) \log \left(1 + \sqrt{2\beta \frac{(1-p) \log n}{Np}} \right) \\ &\quad + \left(1 - p - \sqrt{2\beta \frac{p(1-p) \log n}{N}} \right) \log \left(1 - \sqrt{2\beta \frac{p \log n}{N(1-p)}} \right) \\ &= \left(p + \sqrt{2\beta \frac{p(1-p) \log n}{N}} \right) \left(\sqrt{\frac{2\beta(1-p) \log n}{Np}} - \frac{\beta(1-p) \log n}{Np} + O \left(\frac{(\log n)^{3/2}}{Np} \right) \right) \\ &\quad + \left(1 - p - \sqrt{2\beta \frac{p(1-p) \log n}{N}} \right) \left(-\sqrt{\frac{2\beta p \log n}{N(1-p)}} - \frac{\beta p \log n}{N(1-p)} + O \left(\frac{p \log n}{N} \right)^{3/2} \right) \\ &= \frac{\beta \log n}{N} + O \left(\left(\frac{\log n}{N} \right)^{3/2} \frac{1}{\sqrt{p}} \right) + O \left(\left(\frac{p \log n}{N} \right)^{3/2} \right). \end{aligned}$$

Then by theorem 1 in [2], the third conclusion follows. \square

5.5.2 Proof of Lemma 5.19

Proof of Lemma 5.19. $\mathbb{P}[\mathcal{G}_{s-1}^1] = 1 - o(1)$ follows from standard estimations of binomial variables together with a union bound. To deal with $\mathbb{P}[\mathcal{G}_s^2 \mid \mathcal{F}_{s-1}]$, we first note that \mathcal{G}_s^2 is measurable with respect to \mathbf{E}_s , so we can consider the probability under $\hat{\mathbb{P}}_s$ and conditioning on \mathcal{F}_{s-1} is equivalent to

conditioning on \mathcal{D}_{s-1} . Recall the definition of $\mathcal{A}_{r,s} = \bigcap_{k \in A_{r,s}} \{E_{r,k} < F_k\}$. We start by controlling the conditional probability of $\mathcal{A}_{r,s}$ under realizations of \mathcal{F}_{s-1} that satisfies \mathcal{G}_{s-1}^1 . Under the good event \mathcal{G}_{s-1}^1 , for each $k \in A_{r,s}$, the variable $E_{r,k} = \sum_{j < k} G_{j,k} G_{\pi^*(j),r}$ is dominated by binomial variable $\mathbf{B}(kp + \sqrt{8kp(1-p)\log n}, p)$. Denote $K_0 = kp + \sqrt{8kp(1-p)\log n}$, we have

$$\begin{aligned} & \widehat{\mathbb{P}}_s \left[E_{r,k} < kp^2 + \sqrt{2(1-\varepsilon_n)kp^2(1-p)\log n} \right] \\ & \geq \mathbb{P} \left[\mathbf{B}(K_0, p) < kp^2 + \sqrt{2(1-\varepsilon_n)kp^2(1-p)\log n} \right] \\ & \geq \mathbb{P} \left[\mathbf{B}(K_0, p) < K_0 p + \sqrt{2(1-2\varepsilon_n)K_0 p(1-p)\log n} \right], \end{aligned}$$

which is bounded below by $1 - n^{-1+3\varepsilon_n}$ from Lemma 5.27. Since each event $\{E_{r,k} < F_k\}$ is decreasing, by FKG inequality we get

$$\begin{aligned} \widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}] & \geq \prod_{k \in A_{r,s}} \widehat{\mathbb{P}}_s[E_{r,k} < F_k] \geq (1 - n^{-1+3\varepsilon_n})^{|A_{r,s}|} \geq \exp(n^{3\varepsilon_n}) \\ & = \begin{cases} \exp(-n^{3\eta}), & \text{if } p \leq n^{-0.1}, \\ \exp(-3\sqrt{\log n \log \log n}), & \text{if } n^{-0.1} \leq p \leq 1/(\log n)^3. \end{cases} \end{aligned} \quad (5.43)$$

Note that $(\mathcal{G}_s^2)^c$ is contained in the union of $\mathcal{I}_s^1 = \{|\mathbf{N}_s(r)| \geq sp + \sqrt{sp(1-p)n^{\delta_n}} \text{ for some } r \in \mathbf{R}_{s-1}\}$, $\mathcal{I}_s^2 = \{|\mathbf{N}_s(i) \cap \mathbf{N}_s(j)| \geq 2sp^2 \text{ for some distinct } i, j \in \mathbf{R}_{s-1}\}$ and $\mathcal{D}_s^2 = \{|\mathbf{N}_s(r)| \leq sp - \sqrt{sp(1-p)n^{\delta_n}} \text{ for some } r \in \mathbf{R}_{s-1}\}$. The first two events $\mathcal{I}_s^1, \mathcal{I}_s^2$ are increasing, so

$$\widehat{\mathbb{P}}_s[\mathcal{I}_s^1 \cup \mathcal{I}_s^2 \mid \mathcal{F}_{s-1}] = \widehat{\mathbb{P}}_s[\mathcal{I}_s^1 \cup \mathcal{I}_s^2 \mid \mathcal{D}_{s-1}] \leq \widehat{\mathbb{P}}_s[\mathcal{I}_s^1 \mid \mathcal{D}_{s-1}] + \widehat{\mathbb{P}}_s[\mathcal{I}_s^2 \mid \mathcal{D}_{s-1}] \leq \widehat{\mathbb{P}}_s[\mathcal{I}_s^1] + \widehat{\mathbb{P}}_s[\mathcal{I}_s^2],$$

where the last inequality follows from FKG inequality. By (5.7) and a union bound we see these terms are both $o(1)$. For \mathcal{D}_s^2 , we note that due to independence it holds

$$\widehat{\mathbb{P}}_s[\mathcal{D}_s^2 \mid \mathcal{F}_{s-1}] = \widehat{\mathbb{P}}_s[\mathcal{D}_s^2 \mid \mathcal{D}_{s-1}] = \widehat{\mathbb{P}}_s[\mathcal{D}_s^2 \mid \mathcal{A}_{r,s}] \leq \widehat{\mathbb{P}}_s[\mathcal{D}_s^2] / \widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}].$$

By (5.6) and a union bound we have

$$\widehat{\mathbb{P}}_s[\mathcal{D}_s^2] \leq n \exp(-n^{\delta_n}) = \begin{cases} \exp(\log n - n^{0.1}), & \text{if } p \leq n^{-0.1}, \\ \exp(\log n - \log n \log \log n), & \text{if } n^{-0.1} \leq p \leq 1/(\log n)^3. \end{cases}$$

Combining with the lower bound of $\widehat{\mathbb{P}}_s[\mathcal{A}_{r,s}]$ given by (5.43), we see the conditional probability $\mathbb{P}[\mathcal{D}_s^2 \mid \mathcal{F}_{s-1}]$ is also $o(1)$. This concludes the proof of Lemma 5.19. \square

5.5.3 Proof of Proposition 5.14

Proof of Proposition 5.14. The first item of \mathcal{C} holds with high probability by standard estimation for binomial variables since there is no conditioning for the first ηn vertices.

To show the second item of \mathcal{C} happens with high probability, it suffices to show that for each $\lfloor \eta n \rfloor + 1 \leq s \leq n$, it holds that

$$\mathbb{P}[\forall i \in \mathbf{R}_{s-1}, E_{i,s} \geq sp^2 - \sqrt{10sp^2(1-p)\log n}] = 1 - o(1/n).$$

Define $\mathcal{D}_s^1 = \{|\mathbf{N}_s(r)| \geq sp - \sqrt{sp(1-p)n^{\delta_n}}, \forall r \in \mathbf{R}_{s-1}\}$, where δ_n is defined as before. From the proof of Lemma 5.19 we see that $\mathbb{P}[\mathcal{D}_s^1] = o(1/n)$. Conditioning on any realization of $\mathcal{F}_{s-1} \cup \tilde{\mathcal{N}}_s$ that satisfies \mathcal{D}_s^1 , we see that for each $r \in \mathbf{R}_{s-1}$, $E_{r,s}$ stochastically dominates $\mathbf{B}(sp - \sqrt{1 - sp(1-p)n^{\delta_n}})$. Thus by standard estimates for binomial variables we can show that under \mathcal{D}_s^1 , the conditional probability of $E_{r,s} \geq sp^2 - \sqrt{10sp^2(1-p)\log n}$ for any $r \in \mathbf{R}_{s-1}$ is $1 - o(1/n)$. This completes the proof. \square

5.5.4 Proof of Proposition 5.20

For dense regime, it remains for us to prove proposition 5.20. It is worth mentioning that our proof procedure works for both cases, i.e.

- When $p \leq n^{-0.1}$ and $np^3 \geq n^{-0.1}$, we take $\delta_n = 0.1$ and $\varepsilon_n = \varepsilon$.
- When $n^{-0.1} \leq p \leq 1/(\log n)^3$, we take $\delta_n = (\log \log n + \log \log \log n) / \log n$ and $\varepsilon_n = \log \log n / \sqrt{\log n}$.

Proof of Proposition 5.20. In this proof, we need to show (5.20), which requires us to compute the first moment and the second moment separately.

Computation of the first moment For the first moment, we have,

$$\mathbb{E}[X_s | \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s] = \sum_{i \in \mathbf{R}_{s-1}} \mathbb{P}[E_{i,s} \geq F_s | \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s].$$

For each term in the sum, according to our good condition \mathcal{G}_s^2 , the distribution of $E_{i,s}$ dominates a binomial random variable $\mathbf{B}\left(sp - \sqrt{sp(1-p)n^{\delta_n}}, p\right)$. Let $K'_0 := sp - \sqrt{sp(1-p)n^{\delta_n}}$,

$$\begin{aligned} & \mathbb{P}\left[E_{i,s} \geq F_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right] \geq \mathbb{P}\left[\mathbf{B}(K'_0, p) \geq F_s\right] \\ & \geq \mathbb{P}\left[\mathbf{B}(K'_0, p) \geq K'_0 p + \sqrt{2(1-\varepsilon_n/2)K'_0 p(1-p)\log n}\right] \\ & \stackrel{\text{lem 5.25}}{\geq} 1 - \Phi(\sqrt{2(1-\varepsilon_n/2)\log n}) = O\left(\frac{1}{\sqrt{\log n}} n^{-1+\varepsilon_n/2}\right) \\ & \gg 1/n. \end{aligned}$$

where in the second inequality we use the assumption $p \leq 1/(\log n)^3$. Then it follows that the first moment,

$$\mathbb{E}\left[X_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right] \rightarrow \infty. \quad (5.44)$$

Computation of the second moment In the part, we need to show that,

$$\mathbb{E}\left[X_s^2 \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right] \leq (1 + o(1)) \left(\mathbb{E}\left[X_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right]\right)^2 \quad (5.45)$$

The claim (5.45) is true if we can prove for each $i, j \in \mathbf{R}_{s-1}$ and $i \neq j$,

$$\begin{aligned} & \mathbb{P}\left[E_{i,s} \geq F_s, E_{j,s} \geq F_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right] \\ & \leq (1 + o(1)) \mathbb{P}\left[E_{i,s} \geq F_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right] \mathbb{P}\left[E_{j,s} \geq F_s \mid \mathcal{F}_{s-1}, \tilde{\mathcal{N}}_s\right]. \end{aligned} \quad (5.46)$$

Let $M := |\mathbf{N}_s(i) \cap \mathbf{N}_s(j)|$, $N_1 := |\mathbf{N}_s(i)|$ and $N_2 := |\mathbf{N}_s(j)|$. They all satisfy good event \mathcal{G}_s^2 . We write (5.46) in a simpler form,

$$\mathbb{P}[S_{N_1} \geq F_s, S_{N_2} \geq F_s] \leq (1 + o(1)) \mathbb{P}[S_{N_1} \geq F_s] \mathbb{P}[S_{N_2} \geq F_s]. \quad (5.47)$$

where $S_{N_1} = S_M + S_{N_1-M}$ and $S_{N_2} = S_M + S_{N_2-M}$. The three random variables are independent with distribution,

$$S_{N_i-M} \sim \mathbf{B}(N_i - M, p) \text{ for } i = 1, 2.$$

$$S_M \sim \mathbf{B}(M, p).$$

Before proving (5.47), we remark that the most dangerous case is that $M = 2sp^2 = \theta(np^2)$. This is intuitively clear since when M increases the correlation between S_{N_1} and S_{N_2} grows stronger. To show this, we fix N_1 and N_2 and define the function $F_S(M)$ to be,

$$F_S(M) := \mathbb{P}[S_M + S_{N_1-M} \geq F_s, S_M + S_{N_2-M} \geq F_s]. \quad (5.48)$$

For $F_S(M+1)$, it equals,

$$\begin{aligned}
& \sum_{k=0}^{M+1} \mathbb{P}[S_{M+1} = k] \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k] \\
&= \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_1 = 1] \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k - 1] \\
&\quad + \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_1 = 0] \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k] \tag{5.49} \\
&= \sum_{k=0}^M \mathbb{P}[S_M = k] p \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k - 1] \\
&\quad + \sum_{k=0}^M \mathbb{P}[S_M = k] (1-p) \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k].
\end{aligned}$$

$F_S(M)$ equals,

$$\begin{aligned}
& \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M} \geq F_s - k, S_{N_2-M} \geq F_s - k] \\
&= \sum_{k=0}^M \mathbb{P}[S_M = k] p^2 \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k - 1] \\
&\quad + \sum_{k=0}^M \mathbb{P}[S_M = k] (1-p)^2 \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k] \tag{5.50} \\
&\quad + \sum_{k=0}^M \mathbb{P}[S_M = k] p(1-p) \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k] \\
&\quad + \sum_{k=0}^M \mathbb{P}[S_M = k] p(1-p) \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k - 1]
\end{aligned}$$

Combine (5.49) and (5.50), $F_S(M+1) - F_S(M)$ equals,

$$\begin{aligned}
& p(1-p) \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k - 1] \\
& + p(1-p) \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k] \\
& - p(1-p) \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M-1} \geq F_s - k - 1, S_{N_2-M-1} \geq F_s - k] \\
& - p(1-p) \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M-1} \geq F_s - k, S_{N_2-M-1} \geq F_s - k - 1] \\
& = p(1-p) \sum_{k=0}^M \mathbb{P}[S_M = k] \mathbb{P}[S_{N_1-M-1} = F_s - k - 1, S_{N_2-M-1} = F_s - k - 1] \\
& > 0.
\end{aligned}$$

Since $F_S(M)$ is increasing, it remains for us to prove (5.47) with $M = 2sp^2$ according to the definition of \mathcal{G}_s^2 . Without causing ambiguity we write $M = 2sp^2$ in the following proof. We need to show,

$$\sum_{r=1}^M \mathbb{P}[S_M = r] \frac{\mathbb{P}[S_{N_1-M} \geq F_s - r]}{\mathbb{P}[S_{N_1} \geq F_s]} \frac{\mathbb{P}[S_{N_2-M} \geq F_s - r]}{\mathbb{P}[S_{N_2} \geq F_s]} \leq 1 + o(1). \quad (5.51)$$

To analyze such a summation, we need to consider the following two cases.

- $np^3 \geq (\log n)^2$,
- $n^{-0.1} \leq np^3 \leq (\log n)^2$.

For the first case, by lemma 5.26, we know that,

$$\mathbb{P}[Mp - \sqrt{10Mp(1-p)\log n} \leq S_M \leq Mp + \sqrt{10Mp(1-p)\log n}] = 1 - 1/n^4. \quad (5.52)$$

Let $K_{M,p} := \sqrt{10Mp(1-p)\log n}$. For each $r \in [Mp - K_{M,p}, Mp + K_{M,p}]$, we show that,

$$\frac{\mathbb{P}[S_{N_1-M} \geq F_s - r]}{\mathbb{P}[S_{N_1} \geq F_s]} \leq 1 + o(1). \quad (5.53)$$

To do this, we give a slightly lower bound for the denominator,

$$\begin{aligned}
& \mathbb{P}[S_{N_1} \geq F_s] \\
& \geq \mathbb{P}[S_M \geq Mp - K_{M,p}] \mathbb{P}[S_{N_1-M} \geq F_s - Mp + K_{M,p}] \\
& = (1 - n^{-4}) \mathbb{P}[S_{N_1-M} \geq F_s - Mp + K_{M,p}].
\end{aligned}$$

In this way, we obtain a more tractable form,

$$\frac{\mathbb{P}[S_{N_1-M} \geq F_s - Mp - K_{M,p}]}{\mathbb{P}[S_{N_1-M} \geq F_s - Mp + K_{M,p}]} \leq 1 + o(1). \quad (5.54)$$

For the denominator,

$$\begin{aligned} & \mathbb{P}[S_{N_1-M} \geq F_s - Mp + K_{M,p}] \\ & \geq \mathbb{P}\left[S_{N_1-M} \geq (N_1 - M)p + \sqrt{2(1 - \varepsilon_n)sp^2(1 - p)\log n} + \sqrt{6np^3(1 - p)n^{\delta_n}}\right] \\ & \geq \mathbb{P}\left[S_{N_1-M} \geq (N_1 - M)p + \sqrt{2(1 - \varepsilon_n)(N_1 - M)p(1 - p)\log n} + \sqrt{6np^3(1 - p)n^{\delta_n}}\right. \\ & \quad \left. + \sqrt{2(1 - \varepsilon_n)sp^2(1 - p)\log n} - \sqrt{2(1 - \varepsilon_n)(N_1 - M)p(1 - p)\log n}\right] \\ & \geq \mathbb{P}\left[S_{N_1-M} \geq (N_1 - M)p + \sqrt{2(1 - \varepsilon_n)(N_1 - M)p(1 - p)\log n} + \sqrt{10(N_1 - M)p^2(1 - p)n^{\delta_n}}\right] \\ & \geq \frac{1}{\sqrt{2(1 - \varepsilon_n)\log n} + \sqrt{10pn^{\delta_n}}} \exp((1 - \varepsilon_n)\log n + o(1)), \end{aligned} \quad (5.55)$$

where the last inequality follows from lemma 5.25 and $p \ll 1/(\log n)^3$. We let

$$\begin{aligned} L_{n,p} &:= \sqrt{10np^3(1 - p)n^{\delta_n}}, \\ R_{n,p} &:= (N_1 - M)p + \sqrt{2(1 - \varepsilon_n)(N_1 - M)p(1 - p)\log n}. \end{aligned}$$

Similar to the procedure in (5.55), we have,

$$\mathbb{P}[F_s - Mp - K_{M,p} \leq S_{N_1-M} \leq F_s - Mp + K_{M,p}] \leq \mathbb{P}[R_{n,p} - L_{n,p} \leq S_{N_1-M} \leq R_{n,p} + L_{n,p}] \quad (5.56)$$

We need to show,

$$\mathbb{P}[R_{n,p} - L_{n,p} \leq S_{N_1-M} \leq R_{n,p} + L_{n,p}] = o\left(\frac{1}{\sqrt{\log n}} n^{-1+\varepsilon_n}\right) \quad (5.57)$$

To prove this, for each $k \in (R_{n,p} - L_{n,p}, R_{n,p} + L_{n,p})$, set $N = N_1 - M$, we estimate the quantity $\binom{n}{k} p^k (1-p)^k$ and sum them up. Write $k = Np + M_n + r_n$, where $M_N = \sqrt{2(1 - \varepsilon_n)(N_1 - M)p(1 - p)\log n}$

and r_n be the reminder term with $|r_n| \leq \sqrt{10np^3(1-p)n^{\delta_n}}$.

$$\begin{aligned}
& \binom{N}{k} p^k (1-p)^{N-k} \\
& \leq C \exp \left\{ (k-1/2-N) \log \left(1 - \frac{k}{N} \right) - \frac{1}{2} \log(k) - k \log \left(\frac{k}{N} \right) \right\} p^k (1-p)^{N-k} \\
& = C \exp \left\{ - (k-1/2-N) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{k}{N} \right)^i - \frac{1}{2} \log(Np) - \frac{1}{2} \log \left(\frac{k}{Np} \right) \right. \\
& \quad \left. - k \log \left(\frac{k}{Np} \right) - k \log(p) \right\} p^k (1-p)^{N-k} \\
& = C \exp \left\{ - (k-1/2-N) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{k}{N} \right)^i - \frac{1}{2} \log(Np) + (N-k) \log(1-p) \right. \\
& \quad \left. - \left(k + \frac{1}{2} \right) \log \left(1 + \frac{M_N}{Np} + \frac{r_N}{Np} \right) \right\} \tag{5.58} \\
& = C \exp \left\{ - \sum_{i=1}^{\infty} \frac{1}{i(i+1)} \frac{k^{i+1}}{N^i} - \frac{1}{2} \log(Np) - \frac{1}{2} \frac{M_N^2}{Np} \right. \\
& \quad \left. + N \sum_{i=1}^{\infty} \frac{p^{i+1}}{i(i+1)} + (M_N + r_N) \sum_{i=1}^{\infty} \frac{p^i}{i} + o(1) \right\} \\
& = C \exp \left\{ - \frac{1}{2} \log(Np) - (1 - \varepsilon_n) \log n + o(1) \right\},
\end{aligned}$$

where the $o(1)$ follows from the assumption that $p \ll 1/(\log n)^3$. Thus we have,

$$\begin{aligned}
\mathbb{P}[R_{n,p} - L_{n,p} \leq S_{N_1-M} \leq R_{n,p} + L_{n,p}] &= 2L_{n,p} \max_{k \in (R_{n,p}-L_{n,p}, R_{n,p}+L_{n,p})} \binom{N}{k} p^k (1-p)^{N-k} \\
&= O \left(\frac{n^{\varepsilon_n} \sqrt{pn^{\delta_n}}}{n} \right).
\end{aligned}$$

We have proved (5.57), then it follows (5.53) is true. Combined with (5.52), we immediately know (5.51). The proof for this case is completed.

For the second case $n^{-0.1} \leq np^3 \leq (\log n)^2$, the binomial random variable $\mathbf{B}(M, p)$ behaves in the similar way to $\mathbf{Poi}(Mp)$. By proposition 5.6, we know that,

$$\mathbb{P}[S_M > (\log n)^3] \leq \exp(-C(\log n)^3). \tag{5.59}$$

Recall (5.51), we need to show for each $r \leq (\log n)^3$,

$$\frac{\mathbb{P}[S_{N_1-M} \geq F_s - r]}{\mathbb{P}[S_{N_1} \geq F_s]} \leq 1 + o(1).$$

For simplicity, we prove,

$$\frac{\mathbb{P}[S_{N_1-M} \geq F_s - r]}{\mathbb{P}[S_{N_1-M} \geq F_s]} \leq 1 + o(1). \quad (5.60)$$

Similarly, for the lower bound on the denominator, we have,

$$\begin{aligned} & \mathbb{P}[S_{N_1-M} \geq F_s] \\ & \geq \mathbb{P}\left[S_{N_1-M} \geq N_1 p + \sqrt{2(1-\varepsilon_n)sp^2(1-p)\log n} + \sqrt{6np^3(1-p)n^{\delta_n}}\right] \\ & \geq \mathbb{P}\left[S_{N_1-M} \geq (N_1-M)p + \sqrt{2(1-\varepsilon_n)(N_1-M)p(1-p)\log n} + Mp + \sqrt{6np^3(1-p)\log n}\right. \\ & \quad \left.+ \sqrt{2(1-\varepsilon_n)sp^2(1-p)\log n} - \sqrt{2(1-\varepsilon_n)(N_1-M)p(1-p)\log n}\right] \\ & \geq \frac{1}{\sqrt{2(1-\varepsilon_n)\log n + o(1)}} \exp(-(1-\varepsilon_n)\log n + o(1)). \end{aligned} \quad (5.61)$$

To prove (5.60), we only need to prove the following,

$$\mathbb{P}[F_s - r \leq S_{N_1-M} \leq F_s] = o\left(\frac{1}{\sqrt{\log n}} n^{-1+\varepsilon_n}\right).$$

Apply the procedure in (5.61), the inequality can be reduced to,

$$\mathbb{P}[R_{n,p} - L_{n,p} \leq S_{N_1-M} \leq R_{n,p} + L_{n,p}] = o\left(\frac{1}{\sqrt{\log n}} n^{-1+\varepsilon_n}\right), \quad (5.62)$$

where

$$\begin{aligned} R_{n,p} &:= (N_1 - M)p + \sqrt{2(1-\varepsilon_n)(N_1 - M)p(1-p)\log n}, \\ L_{n,p} &:= \max\left\{(\log n)^3, np^3, \sqrt{6(N_1 - M)p^2(1-p)n^{\delta_n}}\right\}. \end{aligned}$$

To show this, we derive the one-point estimate, i.e. for $k \in (R_{n,p} - L_{n,p}, R_{n,p} + L_{n,p})$, we compute $\mathbb{P}[S_{N_1-M} = k]$ and sum them up. Set $N := N_1 - M$ and let $k = Np + M_n + r_n$, where $M_n = \sqrt{2(1-\varepsilon_n)(N_1 - M)p(1-p)\log n}$ and r_n is the remainder term with $|r_n| \leq L_{n,p}$. For each k , we

have,

$$\begin{aligned}
& \binom{N}{k} p^k (1-p)^{N-k} \\
& \leq C \exp \left\{ (k-1/2-N) \log \left(1 - \frac{k}{N} \right) - \frac{1}{2} \log(k) - k \log \left(\frac{k}{N} \right) \right\} p^k (1-p)^{N-k} \\
& \leq C \exp \left\{ - (k-N) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{k}{N} \right)^i - \frac{1}{2} \log(Np) - \frac{1}{2} \log \left(\frac{k}{Np} \right) \right. \\
& \quad \left. - k \log \left(\frac{k}{Np} \right) - k \log(p) \right\} p^k (1-p)^{N-k} \\
& = C \exp \left\{ - (k-N) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{k}{N} \right)^i - \frac{1}{2} \log(Np) + (N-k) \log(1-p) \right. \\
& \quad \left. - k \log \left(1 + \frac{M_N}{Np} + \frac{r_N}{Np} \right) \right\} \\
& = C \exp \left\{ - (k-N) \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{k}{N} \right)^i - \frac{1}{2} \log(Np) + (N-k) \log(1-p) \right. \\
& \quad \left. - (Np + M_n + r_n) \left(\frac{M_N}{Np} + \frac{r_N}{Np} - 1/2 \left(\frac{M_N}{Np} + \frac{r_N}{Np} \right)^2 + O \left(\left(\frac{M_N}{Np} + \frac{r_N}{Np} \right)^3 \right) \right) \right\} \\
& = C \exp \left\{ - \sum_{i=1}^{\infty} \frac{1}{i(i+1)} \frac{k^{i+1}}{N^i} - \frac{1}{2} \log(Np) - \frac{1}{2} \frac{M_N^2}{Np} \right. \\
& \quad \left. + N \sum_{i=1}^{\infty} \frac{p^{i+1}}{i(i+1)} + (M_N + r_N) \sum_{i=1}^{\infty} \frac{p^i}{i} + o(1) \right\} \\
& = C \exp \left\{ - \frac{1}{2} \log(Np) - (1 - \varepsilon_n) \log n + o(1) \right\},
\end{aligned} \tag{5.63}$$

where the $o(1)$ comes from the definition of M_n and r_n , and the range of p and δ_n . Then summing over all possible k yields,

$$\begin{aligned}
& \mathbb{P}[R_{n,p} - L_{n,p} \leq S_{N_1-M} \leq R_{n,p} + L_{n,p}] \\
& \leq 2L_{n,p} \max_{k \in (R_{n,p} - L_{n,p}, R_{n,p} + L_{n,p})} \binom{N}{k} p^k (1-p)^{N-k} \\
& \leq 2L_{n,p} \frac{C}{\sqrt{Np}} n^{-1+\varepsilon_n} = o \left(\frac{1}{\sqrt{\log n}} n^{-1+\varepsilon_n} \right),
\end{aligned} \tag{5.64}$$

where the last equality follows from the definition of $L_{n,p}$. This proves (5.62), which implies that (5.60) is true. Combining (5.51) and (5.59), we prove (5.47), thus finishing the proof of the second case. \square

References

- [1] N. Alon and J. H. Spencer. The probabilistic method. *Wiley Publishing*, 4th edition, 2016.
- [2] Richard Arratia and Louis Gordon. Tutorial on large deviations for the binomial distribution. *Bulletin of mathematical biology*, 51(1):125131, 1989.
- [3] A. Berg, T. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondences. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 26–33 vol. 1, 2005.
- [4] B. Barak, C.-N. Chou, Z. Lei, T. Schramm, and Y. Sheng. (Nearly) efficient algorithms for the graph matching problem on correlated random graphs. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [5] Rainer E Burkard, Eranda Cela, Panos M Pardalos, and Leonidas S Pitsoulis. The quadratic assignment problem. In *Handbook of combinatorial optimization*, pages 17131809. Springer, 1998.
- [6] A. S. Bandeira, A. Perry, and A. S. Wein. Notes on computational-to-statistical gaps: predictions using statistical physics. *Port. Math.*, 75(2):159–186, 2018.
- [7] D. Cullina and N. Kiyavash. Exact alignment recovery for correlated Erdos-Rényi graphs. Preprint, arXiv:1711.06783.
- [8] D. Cullina and N. Kiyavash. Improved achievability and converse bounds for erdos-renyi graph matching. In *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, SIGMETRICS '16, page 6372, New York, NY, USA, 2016. Association for Computing Machinery.
- [9] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006.
- [10] J. Ding and H. Du. Detection threshold for correlated erdos-renyi graphs via densest subgraph. Preprint, arXiv:2203.14573.
- [11] J. Ding and H. Du. Matching recovery threshold for correlated random graphs. Preprint, arXiv:2205.14650.

- [12] J. Ding, H. Du, and S. Gong. A polynomial-time approximation scheme for the maximal overlap of two independent Erdos-Rényi graphs. Preprint, arXiv:2210.07823.
- [13] J. Ding, Z. Ma, Y. Wu, and J. Xu. Efficient random graph matching via degree profiles. *Probab. Theory Related Fields*, 179(1-2):29–115, 2021.
- [14] D. Gamarnik. The overlap gap property: A topological barrier to optimizing over random structures. *Proceedings of the National Academy of Sciences*, 118(41):e2108492118, 2021.
- [15] L. Ganassali, L. Massoulié, and M. Lelarge. Correlation detection in trees for planted graph alignment To appear in *Ann. Appl. Probab.*
- [16] D. Gamarnik, E. C. Kzldag, W. Perkins and C. Xu. Geometric Barriers for Stable and Online Algorithms for Discrepancy Minimization. Preprint, arXiv:2302.06485.
- [17] D. Gamarnik and M. Sudan. Limits of local algorithms over sparse random graphs. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 369376. ACM, 2014.
- [18] D. Gamarnik and I. Zadik. The landscape of the planted clique problem: Dense subgraphs and the overlap gap property. Preprint, arXiv:1904.07174.
- [19] G. Hall and L. Massoulié. Partial recovery in the graph alignment problem. Preprint, arXiv:2007.00533.
- [20] A. Haghighi, A. Ng, and C. Manning. Robust textual inference via graph matching. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 387–394, Vancouver, British Columbia, Canada, Oct 2005.
- [21] B. Huang and M. Sellke, Tight Lipschitz Hardness for optimizing Mean Field Spin Glasses, *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, Denver, CO, USA, 2022, pp. 312-322.
- [22] Mohammed El-Kebir, Jaap Heringa, and Gunnar W Klau. Lagrangian relaxation applied to sparse global network alignment. In *Pattern Recognition in Bioinformatics*. Springer, 2011, pp. 225236.

- [23] Oleksii Kuchaiev and Nataa Prulj. Integrative network alignment reveals large regions of global network similarity in yeast and human. In *Bioinformatics* 27.10 (2011), pp. 13901396.
- [24] G. Kalai and S. Safra. Perspectives from mathematics, computer science, and economics. *Computational complexity and statistical physics*, page 25, 2006.
- [25] Konstantin Makarychev, Rajsekar Manokaran, and Maxim Sviridenko. Maximum quadratic assignment problem: Reduction from maximum label cover and lp-based approximation algorithm. *Automata, Languages and Programming*, pages 594604, 2010.
- [26] A. Montanari. Optimization of the sherrington-kirkpatrick hamiltonian. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1417–1433, 2019.
- [27] P. Manurangsi, A. Rubinstein, and T. Schramm. The Strongish Planted Clique Hypothesis and Its Consequences. In J. R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:21, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.
- [28] C. Mao, M. Rudelson, and K. Tikhomirov. Exact matching of random graphs with constant correlation. To appear in *Probab. Theory and Related Fields*.
- [29] C. Mao, Y. Wu, J. Xu and S. H. Yu, Random graph matching at Otter’s threshold via counting chandeliers. preprint, arXiv:2209.12313.
- [30] A. Narayanan and V. Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, 2008.
- [31] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *2009 30th IEEE Symposium on Security and Privacy*, pages 173–187, 2009.
- [32] Panos M. Pardalos, Franz Rendl, and Henry Wolkowicz. The quadratic assignment problem: A survey and recent developments. In *In Proceedings of the DIMACS Workshop on Quadratic Assignment Problems*, volume 16 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science, pages 142. American Mathematical Society, 1994.
- [33] P. Raghavendra, T. Schramm, and D. Steurer. High dimensional estimation via sum-of-squares proofs. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 3389–3423. World Sci. Publ., Hackensack, NJ, 2018.

- [34] M. Rahman and B. Virág. Local algorithms for independent sets are half-optimal. *Ann. Probab.*, 45(3):1543–1577, 2017.
- [35] Eric V. Slud. Distribution inequalities for the binomial law. *Ann. Probab.*, 5(3):404–412, 1977.
- [36] E. Subag. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.*, 74(5):1021–1044, 2021.
- [37] R. Singh, J. Xu, and B. Berger. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proceedings of the National Academy of Sciences of the United States of America*, 105:12763–8, 10 2008.
- [38] J. T. Vogelstein, J. M. Conroy, V. Lyzinski, L. J. Podrazik, S. G. Kratzer, E. T. Harley, D. E. Fishkind, R. J. Vogelstein, and C. E. Priebe. Fast approximate quadratic programming for graph matching. *PLOS ONE*, 10(4):1–17, 04 2015.
- [39] A. S. Wein. Optimal low-degree hardness of maximum independent set. *Mathematical Statistics and Learning*, 2022.
- [40] Y. Wu and J. Xu. *Statistical Problems with Planted Structures: Information-Theoretical and Computational Limits*, page 383–424. Cambridge University Press, 2021.
- [41] Y. Wu, J. Xu and S. H. Yu, Testing correlation of unlabeled random graphs. To appear in *Ann. Appl. Probab.*
- [42] Y. Wu, J. Xu and S. H. Yu, Settling the Sharp Reconstruction Thresholds of Random Graph Matching, *2021 IEEE International Symposium on Information Theory (ISIT)*, Melbourne, Australia, 2021, pp. 2714–2719.
- [43] L. Zdeborová and F. Krzakala. Statistical physics of inference: thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.

6 Percolation phase transition of Brownian loop soup on metric graphs

In this section, we show that the critical threshold for the percolation of the Brownian loop soup on a large class of transient metric graphs (including quasi-transitive graphs such as \mathbb{Z}^d , $d \geq 3$) is $1/2$. This section is based on a joint work with Yinshan Chang and Xinyi Li.

6.1 Introduction and the main result

The model of Brownian loop soup was first introduced by Lawler and Werner in the seminal paper [7] as a Poissonian collection of loops whose law is based on that of the Brownian motion. Its random walk analog, the random walk loop soup, was introduced by Lawler and Trujillo Ferreras in [6]. Loop soups are intimately related to various objects of interest in probability and statistical physics, in particular via the isomorphism theorem discovered by Le Jan [8] linking the loop soup of intensity¹ $1/2$ to the Gaussian free field.

The percolation of loop soups was already considered by Lawler and Werner in [7] and then by Sheffield and Werner in [12] in the setting of the two-dimensional Brownian loop soup. The latter paper, among other results, identified the value of the critical intensity as $1/2$. In the discrete setting, Lupu in [11] also identified the critical threshold in the case of the upper half plane. Subsequently, the works [9, 2, 1] considered loop percolation on \mathbb{Z}^d for $d \geq 3$ and established various results regarding the phase transition in percolative properties.

In [10], Lupu considered the Brownian loop soup on the so-called “metric graphs” (also referred to as the “cable system” or “cable graphs”), a notion that corresponds to the extension of discrete graphs to a continuous metric space in which each edge of the graph have a “length” and Markov chains are embedded in Brownian motions moving continuously along edges. This particular model interpolates between the discrete and the continuum, on which the power of the link to Gaussian free field is maximized, yielding exact formulas (in particular the two-point function from [10]; see Proposition 6.5 below for more details) that allow the authors of [3] to conclude that the critical threshold for the percolation of Brownian loop soup on a large class of transient metric graphs

¹Note that in the early literature there is an inconsistency of a multiplicative factor of 2 in the intensity parameter from the definition of loop soups; see e.g. [11] for a detailed discussion on this issue.

(including the metric version of \mathbb{Z}^d , $d \geq 3$ and regular trees) is greater or equal to $1/2$. It remained an open question whether this threshold is exactly equal to $1/2$. (In contrast, the critical threshold for the discrete loop percolation does not equal to $1/2$ in general, see e.g. [2, Theorem 1.3] where it is shown that the threshold for the discrete loop percolation on \mathbb{Z}^d tends to infinity as $d \rightarrow \infty$.)

In this short note, we give a positive answer to this question on a sufficiently general class of metric graphs by a simple application of the Russo's formula and the two-point function discovered in [10].

We now state our main result. Given a metric graph G , we denote by \mathbb{P}_α for the law of a Brownian loop soup on G with intensity $\alpha > 0$ and by $x_o \longleftrightarrow \infty$ the event that $x_o \in G$ is in an infinite cluster formed by the loop soup.

Theorem 6.1. *For any quasi-transitive transient metric graph G and any $x_o \in G$, it holds that*

$$\mathbb{P}_\alpha[x_o \longleftrightarrow \infty] > 0 \iff \alpha > 1/2. \quad (6.1)$$

In other words, the critical threshold for the loop percolation on G is $1/2$.

Remark 6.2. In fact, our proof works for more general metric graphs, see Remark 6.9 for discussions on sufficient conditions for our result to hold.

Acknowledgments: The authors acknowledge the support of National Key R&D Program of China (No. 2021YFA1002700 and No. 2020YFA0712900). YC acknowledges the support of NSFC (No. 11701395). HD is partially supported by the elite undergraduate training program of School of Mathematical Science at Peking University. XL acknowledges the support of NSFC (No. 12071012).

6.2 Preliminaries

In this section, we briefly introduce metric graphs as well as the associated Brownian loop soup and discuss a few classical preliminary facts that will be useful to the proof. For a more detailed introduction of metric graphs and related objects, see e.g. Section 2 of [10].

We start with metric graphs.

Definition 6.3 (Metric graphs). *Let $G^{\text{skeleton}} = (V, E, \lambda)$ be an unoriented finite or countably-infinite weighted connected graph of finite degrees such that $\lambda_{x,y} > 0$ for any $x, y \in V$ and $w(x) :=$*

$\sum_{y \sim x} \lambda_{x,y} < \infty$ for any $x \in V$. Then the metric graph² G associated with G^{skeleton} is the metric space where each edge $e \in E$ is regarded as an interval of length $(2\lambda_e)^{-1}$, referred to as the **metric graph** G with skeleton G^{skeleton} .

If in addition there is a finite subset $V_o \subset V$ such that for any $x \in V$ there is an automorphism of G^{skeleton} mapping x to some $x_o \in V_o$ (this automorphism should also preserve weights), we say that G is **quasi-transitive**. Examples of quasi-transitive metric graphs include the metric version of periodic lattices (in particular, \mathbb{Z}^d) and regular trees with a fixed set of possible choices of edge weights.

We now turn to the Brownian motion and Brownian loop soup on metric graphs.

Given a metric graph G , there is an associated canonical diffusion process, called the Brownian motion $(B_t^G)_{t \geq 0}$ on G . We write $(l_y)_{y \in G}$, Q , and E for the corresponding local time process, the probability law and expectation respectively. If G is transient, it is possible to define the Green's function $G(\cdot, \cdot) : G \times G \rightarrow \mathbb{R}^+$ associated with B^G as

$$G(x, y) = E_x[l_y], \forall x, y \in G,$$

i.e., the expected value of the local time at y of a Brownian motion B^G on G starting from x . When $x, y \in V$, $G(x, y)$ coincides with the Green's function associated with the Markov jump process on G^{skeleton} . For any open subset K of G , it is also possible to define the killed Green's function $G_K(x, y)$ for $x, y \in K$, as the expected local time at y before the Brownian motion started from x exits K for the first time.

Definition 6.4 (The Brownian loop soup). *Given a metric graph G , we can endow a σ -finite measure μ on the spaces of (rooted) loops on G (i.e. continuous paths $l : [0, T] \rightarrow G$ such that $l(0) = l(T)$) defined via*

$$\mu = \int_{x \in G} \int_0^\infty Q_{x,x}^t p_t(x, x) \frac{dt}{t} dx$$

where we denote by $Q_{x,x}^t$ and $p_t(x, x)$ the bridge probability measure and transition kernel of B^G from x to x of duration t respectively. A Brownian loop soup with intensity $\alpha > 0$ on the metric graph G , denoted by BLS_α is a Poisson point process on the set of loops in G with intensity $\alpha\mu$.

²Although in the literature metric graphs are usually denoted with a tilde as in " \tilde{G} " in contrast to its skeleton, in this note we do not follow this convention as we will be almost exclusively working on metric graphs.

A configuration ω of BLS_α , can be viewed as an open subset of the metric graph, which is the union of ranges of (random) loops in the point process. Denote $\mathbb{P}_\alpha = \mathbb{P}_\alpha^G$ for the probability measure of configurations sampled from BLS_α .

We now turn to loop percolation. For two Borel subsets A, B of G and a configuration $\omega \sim \text{BLS}_\alpha$, we say $A \longleftrightarrow B$ if A, B are connected by a path entirely lying in ω . With slight abuse of notation we also write events like $x \longleftrightarrow \cdot$ as a shorthand for $\{x\} \longleftrightarrow \cdot$. For any Borel subset $K \subset G$, the event such that $A \cap K$ is connected to $B \cap K$ via clusters of loops of ω **entirely** lying in K is denoted by $A \xleftrightarrow{K} B$. Given $x_o \in G$, write $\{x_o \longleftrightarrow \infty\}$ for the event that x_o lies in an infinite cluster of the loop soup configuration ω .

The following two-point function estimate is a paraphrase of Proposition 5.2 of [10].

Proposition 6.5 (Killed two-point function). *For any open subset $K \subset G$ and $x, y \in K$, it holds that*

$$\mathbb{P}_{1/2}[x \xleftrightarrow{K} y] = \frac{2}{\pi} \arcsin \frac{G_K(x, y)}{\sqrt{G_K(x, x)G_K(y, y)}}. \quad (6.2)$$

For any $\alpha > 1/2$, we note that a configuration $\omega \sim \text{BLS}_\alpha$ has the same distribution of the superposition of ω_1 and ω_2 , where ω_1, ω_2 are independent configurations sampled from $\mathbb{P}_{1/2}$ and $\mathbb{P}_{\alpha-1/2}$, respectively. By the quasi-transitivity of G , it is clear that ω_2 stochastically dominates a “Bernoulli 2-bond percolation model” on G where we independently open any 2-bond, i.e., pair of **neighboring** edges with probability $p = p(\alpha - 1/2, G) > 0$. Motivated by this, we denote $\overline{\mathbb{P}}_\varepsilon$ as the law of superposing of two independent configurations $\omega_1 \sim \mathbb{P}_{1/2}$ and ω_ε^b , that of a Bernoulli 2-bond percolation with parameter $\varepsilon \in [0, 1)$, whose law we denote by P_ε .

Let $E^2(G)$ stand for the set of 2-bonds in G . For an increasing event A and a configuration ω , we say a 2-bond $ef \in E^2(G)$ is pivotal for A in ω , if $\omega \cup \{e, f\} \in A$ but $\omega \setminus \{e, f\} \notin A$. We have the following Russo’s formula:

Proposition 6.6 (Russo’s formula). *For any $\varepsilon \geq 0$, and any increasing event A which depends on a finite range of edges, it holds that as $\delta \downarrow 0$,*

$$\overline{\mathbb{P}}_{\varepsilon+\delta}[A] - \overline{\mathbb{P}}_\varepsilon[A] = \delta \cdot \overline{\mathbb{E}}_\varepsilon[\#\{ef \in E^2(G) : ef \text{ is pivotal for } A\}] + O(\delta^2). \quad (6.3)$$

We omit the proof as it follows from rather standard arguments.

Finally, we turn to the FKG inequality for the new measure $\overline{\mathbb{P}}_\varepsilon$, which is a direct consequence of the FKG inequality for general Poisson processes (see Lemma 2.1 of [5]) and the observation that

when looking at percolative properties, one can effectively regard $\omega \sim \bar{\mathbb{P}}_\varepsilon$ as a Poisson process in $\mathcal{L}_G \cup E^2(G)$.

Proposition 6.7 (FKG inequality). *For $\varepsilon \in [0, 1)$ and any increasing events A, B ,*

$$\bar{\mathbb{P}}_\varepsilon[A \cap B] \geq \bar{\mathbb{P}}_\varepsilon[A] \bar{\mathbb{P}}_\varepsilon[B].$$

6.3 Proof of the main result

Without loss of generality we assume $x_o \in V$. Let $d(\cdot, \cdot)$ be the **discrete** graph metric on G^{skeleton} . We write

$$B_n = \{x \in V : d(x_o, x) \leq n\} \quad \text{and} \quad \partial^s B_n = \{x \in V : d(x_o, x) = n\}.$$

We then write

$$f_n(\varepsilon) := \bar{\mathbb{P}}_\varepsilon[x_o \longleftrightarrow \partial^s B_n].$$

The crux of the proof is the following differential inequality, which is remotely inspired by the argument in [4]

Proposition 6.8. *There exist constants $c, C > 0$ (depending on G and x_o only), such that for any sufficiently large n , it holds*

$$f'_{n,+}(\varepsilon) \geq c(1 - C f_n(\varepsilon)), \quad \forall 0 \leq \varepsilon < 1, \quad (6.4)$$

where $f'_{n,+}$ stands for the right derivative of f_n .

Proof. From Proposition 6.6, it suffices to show that there exists $c > 0$ such that for any $\varepsilon \geq 0$,

$$\bar{\mathbb{E}}_\varepsilon[\#\{ef \text{ is pivotal for } x_o \longleftrightarrow \partial^s B_n\}] \geq c(1 - C f_n(\varepsilon)).$$

For any configuration ω sampled from $\bar{\mathbb{P}}_\varepsilon$, denote ω_{B_n} as its restriction in B_n . We consider the union of connected components of ω_{B_n} intersecting $\partial^s B_n$, and denote its closure by $C_n = C_n(\omega)$. Define K_n as the component of $B_n \setminus C_n$ containing x_o (if $x_o \in C_n$, then set $K_n = \emptyset$). By exploring from $\partial^s B_n$ inwards, we see that conditioned on the realization of C_n , the configuration in K_n has the law of the superposition of the loop soup of intensity $1/2$ in K_n and the Bernoulli 2-bond percolation of parameter ε on $K_n^s := V(G) \cap K_n$, independent with the configurations in C_n .

Write $C_n^s := V(G) \cap C_n$. Let

$$E(K_n, C_n) = \{ef = (x, y)(y, z) \in E^2(G) : x, y \in K_n^s, z \in C_n^s\}$$

be the boundary 2-bonds of K_n and let

$$\partial_i K_n := \{x \in K_n^s : d(x, C_n^s) = 1\} \quad \text{and} \quad \partial_i^2 K_n := \{x \in K_n^s : d(x, C_n^s) = 2\}$$

stand for the inner vertex boundary and 2-inner vertex boundary of K_n respectively. Write

$$\mathcal{A}_n := \{d(x_o, C_n^s) > 2\}.$$

Note that \mathcal{A}_n is measurable with respect to the realization of C_n .

We now make the following observation: conditioned on any realization of C_n such that \mathcal{A}_n holds, for any $ef = (x, y)(y, z) \in E(K_n, C_n)$, ef is pivotal for $x_o \longleftrightarrow \partial^s B_n$ if $x_o \xleftrightarrow{K_n} x$. As a result we have

$$\begin{aligned} & \mathbb{E}_\varepsilon[\#\{ef \in E(K_n, C_n); ef \text{ is pivotal for } x_o \longleftrightarrow \partial^s B_n\}] \\ & \geq \mathbb{E}_\varepsilon\left[\sum_{ef \in E(K_n, C_n)} \mathbb{P}_\varepsilon[ef \text{ is pivotal for } x_o \longleftrightarrow \partial^s B_n \mid C_n]\right] \\ & \geq \mathbb{P}_\varepsilon[\mathcal{A}_n] \cdot \mathbb{E}_\varepsilon\left[\sum_{e=(x,y)(y,z) \in E(K_n, C_n), x \in \partial_i^2 K_n} \mathbb{P}_\varepsilon[x_o \xleftrightarrow{K_n} x \mid C_n] \mid \mathcal{A}_n\right] \\ & \geq \mathbb{P}_\varepsilon[\mathcal{A}_n] \cdot \mathbb{E}_\varepsilon\left[\sum_{x \in \partial_i^2 K_n} \mathbb{P}_{1/2}[x_o \xleftrightarrow{K_n} x \mid C_n] \mid \mathcal{A}_n\right]. \end{aligned} \tag{6.5}$$

To estimate the first term in (6.5), we define $\mathcal{O}(x_o)$ for the event that all edges incident to x_o and graph neighbors of x_o are covered by a single loop in the 1/2-BLS. Also we note that \mathcal{A}_n^c is equivalent to the event that one of the neighbors or 2-neighbors of x_o in $V(G)$ is connected to $\partial^s B_n$. Then it is clear that both \mathcal{A}_n and $\mathcal{O}(x_o)$ are increasing, and $\mathcal{A}_n^c \cap \mathcal{O}(x_o) \subset \{x_o \longleftrightarrow \partial^s B_n\}$. By FKG inequality (see Proposition 6.7), we get

$$\mathbb{P}_\varepsilon[\mathcal{A}_n] = 1 - \mathbb{P}_\varepsilon[\mathcal{A}_n^c] \geq 1 - \frac{\mathbb{P}_\varepsilon[\mathcal{A}_n^c \cap \mathcal{O}(x_o)]}{\mathbb{P}_{1/2}[\mathcal{O}(x_o)]} \geq 1 - C \mathbb{P}_\varepsilon[x_o \longleftrightarrow \partial^s B_n] = 1 - C f_n(\varepsilon),$$

where $C = C(G) = \sup_{x_o \in V(G)} (\mathbb{P}_{1/2}[\mathcal{O}(x_o)])^{-1}$ is a positive constant depending on G only. For

the second term, from (6.2) we see that

$$\begin{aligned} \sum_{x \in \partial_i^2 K_n} \mathbb{P}_{1/2}[x_o \xleftrightarrow{K_n} x \mid C_n] &= \sum_{x \in \partial_i^2 K_n} \frac{2}{\pi} \arcsin \frac{G_{K_n}(x_o, x)}{\sqrt{G_{K_n}(x_o, x_o) G_{K_n}(x, x)}} \\ &\geq c_1 \sum_{x \in \partial_i^2 K_n} \frac{G_{K_n}(x_o, x)}{\sqrt{G_K(x, x)}}, \end{aligned} \quad (6.6)$$

where $c_1 = 2(\pi \sqrt{G(x_o, x_o)})^{-1}$ depends only on G and x_o , and we have used the fact that $\arcsin x \geq x$ for $x \in (0, 1]$ and $G_{K_n}(x_o, x_o) \leq G(x_o, x_o)$ for any K_n . Furthermore, we have for each $x \in \partial_i^2 K_n$,

$$\begin{aligned} \frac{G_{K_n}(x_o, x)}{\sqrt{G_{K_n}(x, x)}} &= \sqrt{G_{K_n}(x, x)} \cdot \frac{G_{K_n}(x_o, x)}{G_{K_n}(x, x)} \\ &= \sqrt{G_{K_n}(x, x)} \cdot Q_{x_o}[B^G \text{ hits } x \text{ before exiting } K_n]. \end{aligned} \quad (6.7)$$

In addition, $x \in \partial_i^2 K_n$ implies $G_{K_n}(x, x) \geq c_2$ for some constant $c_2 > 0$ depending only on G .

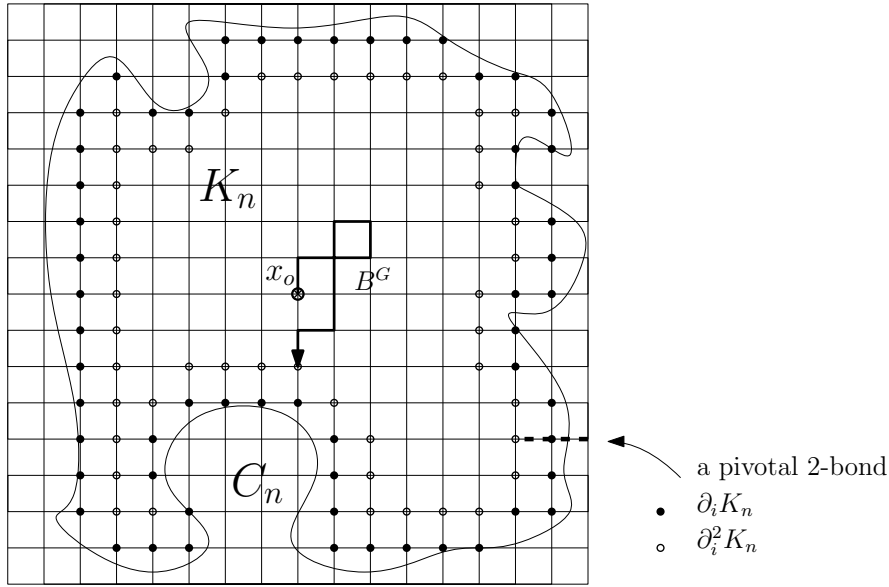


Figure 6.1: A possible realization of K_n , C_n , $\partial_i K_n$, $\partial_i^2 K_n$ and an exemplary pivotal 2-bond. Note that in this figure B_n and C_n are only partially depicted.

Combining (6.6) with (6.7), we see that the second term in (6.5) is bounded from below by

$$c_1 \sqrt{c_2} \sum_{x \in \partial_i^2 K_n} Q_{x_o}[B^G \text{ hits } x \text{ before leaving } K_n].$$

Note that under the event \mathcal{A}_n , the sum in the above formula is trivially bounded from below by 1, since a path of the Brownian motion B^G on G starting from x_o almost surely hits at least one vertex $x \in \partial_i^2 K_n$ before exiting K_n . Therefore, the proof is completed by taking $c = c_1 \sqrt{c_2}$. \square

It is worth noting that there is a difficulty in the argument above if we only consider $\partial_i K_n$, the interior graph boundary of K_n (instead of its 2-inner boundary) in summing the two point functions. Indeed, for some $x \in \partial_i K_n$, it could happen that x is very close to C_n in the metric sense, and in this case one may not bound $G_{K_n}(x, x)$ from below by any universal constant.

Proof of Theorem 6.1. First, it is easy to see that quasi-transitive transient weighted graphs satisfy the (Cap) condition in [3], hence by Corollary 3.6, *ibid.*, $f(0) = 0$, and the critical threshold for loop soup percolation on G is greater or equal to $1/2$.

The main claim then follows from a standard and classical argument. It suffices to show that for any $\varepsilon > 0$,

$$f(\varepsilon) := \overline{\mathbb{P}}_\varepsilon[x_o \longleftrightarrow \infty] > 0.$$

For any $\varepsilon \in (0, 1)$, if $f(\varepsilon) \geq 1/(2C)$, where C is the constant from Proposition 6.8, then the claim naively follows; otherwise, one can find $N_0 = N_0(\varepsilon)$ such that $f_n(\varepsilon) < 1/(2C)$ for all $n \geq N_0$. By Proposition 6.8 and the monotonicity of $f_n(\cdot)$, uniformly for all $\delta \in (0, \varepsilon)$ and $n > N_0$, one has $f'_{n,+}(\delta) \geq c(1 - Cf_n(\varepsilon)) > c/2$. From this we conclude that $f_n(\varepsilon) > c\varepsilon/2$ for any $n \geq N_0$. Combining the two cases, one has $f(\varepsilon) > 0$ for any $\varepsilon > 0$. \square

Remark 6.9. 1) Our proof does not essentially rely on the quasi-transitivity of G except at places where we require a uniform constant depending on the graphs G . In fact, it applies to all transient weighted graphs satisfying the (Cap) condition from [3] with uniformly bounded degrees and edge weights uniformly bounded from above (but not necessarily from below – since one may always break a “long” edge into shorter pieces, which produces a new skeleton graph but keep the metric space (of the metric graph) and the percolation threshold unchanged).

2) Another direction of generalization is to take killing (which corresponds to massive loop soups and or loop soups on a metric graph with boundary) into consideration. If the killing is mild, i.e.,

$$h_{\text{kill}}(x) := Q_x[B^G \text{ is killed}] < 1, \forall x \in G,$$

(note that this is a necessary condition for (Cap) condition from [3] to hold), then our arguments still work with little modification. It is then very natural to ask the following question: what is the

sufficient and necessary condition one should pose on a metric graph for the critical threshold of loop percolation to be exactly $1/2$?

References

- [1] Yinshan Chang. Supercritical loop percolation on \mathbb{Z}^d for $d \geq 3$. *Stochastic Process. Appl.*, 127(10):3159-3186, 2017.
- [2] Yinshan Chang and Artem Sapozhnikov. Phase transition in loop percolation. *Probab. Theory Relat. Fields*, 164(3-4):979-1025, 2016.
- [3] Alexander Drewitz, Alexis Prévost and Pierre-François Rodriguez. Cluster capacity functionals and isomorphism theorems for Gaussian free fields. *Probab. Theory Relat. Fields*, 183(1-2):255-313, 2022.
- [4] Hugo Duminil-Copin and Vincent Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Comm. Math. Phys.*, 343(2):725-745, 2016.
- [5] Svante Janson. Bounds on the distributions of extremal values of a scanning process. *Stochastic Process. Appl.*, 18(2):313-328, 1984.
- [6] Gregory F. Lawler and José A. Trujillo Ferreras. Random walk loop soup. *Trans. Amer. Math. Soc.*, 359(2):767-787, 2007.
- [7] Gregory F. Lawler and Wendelin Werner. The Brownian loop soup. *Probab. Theory Related Fields*, 128(4):565-588, 2004.
- [8] Yves Le Jan. Markov paths, loops and fields, in: *Lecture Notes in Mathematics*, vol. 2026, Springer, 2011.
- [9] Yves Le Jan and Sophie Lemaire. Markovian loop clusters on graphs. *Illinois J. Math.* 57(2):525-558, 2013.
- [10] Titus Lupu. From loop clusters and random interlacements to the free field. *Ann. Probab.* 44(3):2117-2146, 2016.
- [11] Titus Lupu. Loop percolation on discrete half-plane, *Electron. Commun. Probab.*, 21(30):1-9, 2016.

- [12] S. Sheffield and W. Werner. Conformal loop ensembles: the Markovian characterization and the loop-soup construction. *Ann. of Math.*, 176(**3**):1827-1917, 2012.

7 Sharp asymptotics for arm probabilities in critical planar percolation

In this section, we consider critical planar site percolation on the triangular lattice and derive sharp estimates on the asymptotics of the probability of half-plane j -arm events for $j \geq 1$ and planar (polychromatic) j -arm events for $j > 1$. These estimates greatly improve previous results and in particular answer (a large part of) a question of Schramm (*ICM Proc.*, 2006). This section is based on a joint work with Yifan Gao, Xinyi Li and Zijie Zhuang.

7.1 Introduction

Percolation is without doubt one of the most studied statistical mechanics models in probability. As an ideal playground for the study of phase transitions and criticality, it has received considerable attention from probabilists and statistic physicists in the past 60+ years. Despite its simple setup, it is the source of many fascinating yet difficult mathematical problems, with some already well answered and many more still very far from being solved. Starting from the beginning of 21st century, there have been a lot of breakthroughs in the study of a particular case of this model, namely the critical planar percolation on the triangular lattice. In the ground-breaking work by Smirnov [24], it is shown that the both crossing probabilities and the exploration process have conformally invariant scaling limits that can be given through Cardy's formula and described as a (chordal) Schramm-Loewner evolution with parameter 6 (SLE_6) respectively. Later, in [5] Camia and Newman give the characterization of the full scaling limit, via the collection of non-self-crossing loops that is known now as the conformal loop ensemble (CLE). To keep the introduction concise, we refer readers to the classical book [9] and a recent survey [6] for more on the history and some recent progresses of this model.

Critical exponents are central notions in understanding the behavior of critical models. In the case of percolation, many of them can be derived from the so-called arm exponents, i.e. exponents in the power-law decay of the probability of arm events (as the mesh size tends to 0), which we will introduce shortly. For critical planar percolation on the triangular lattice, Smirnov and Werner calculate the precise values of the half-plane arm exponents and (polychromatic) j -arm exponents with $j > 1$ in [25], using Kesten's scaling relation from [12], Cardy's formula as well as relations

between interfaces and SLE_6 . Roughly in the same time, Lawler, Schramm and Werner obtain the one-arm exponent in the plane in [16], also via SLE-related calculations. We will briefly recall their results in Claim (3) of Lemma 7.11 and (7.10) respectively.

However, these asymptotics (with an $o(1)$ in the exponent) are not completely satisfactory. One particular reason is that in dealing with scaling limits involving microscopic quantities from critical planar percolation, such as the pivotal, cluster or interface measures from [8] (see in particular Theorems 4.3, 5.1 and 5.5, *ibid.*) and the natural parametrization of the interface (see Theorem 1.4 in [10]), instead of renormalizing with an explicit factor η^α where η is the mesh size and α the corresponding exponent, one normalizes by something implicit, namely the asymptotic arm probability (which corresponds to a_j in this paper); see also Remark 4.10 in [8] for more on this issue.

Thus, it is natural to ask if there are precise estimates of the arm events, just as what have been obtained for other critical models, e.g., sharp asymptotics of the one-point function of the loop-erased random walk (LERW); see [11], [17] and [14] for the two-, three- and four-dimensional cases respectively.

In fact, much before the scaling limit results mentioned in the paragraph above are obtained, Schramm already asks the question in the proceedings of ICM 2006 if it is possible to improve the estimates. In Problem 3.1 of [22], he points out that “it would be especially nice to obtain estimates that are sharp up to multiplicative constants”.

In a few special cases³, both the value and up-to-constants estimates can be obtained without relating to SLE’s; see Lemma 5 of [13], Theorem 24 of [19] or the first exercise sheet of [27]. Improvements for other arm asymptotics are much more difficult. Mendelson, Nachmias and Watson obtain in [18] a power-law rate for the Cardy’s formula⁴ which yields a slight improvement of half-plane one-arm asymptotics (see (7.2)), although still not as strong as up-to-constants estimates.

In this paper, we answer a large part⁵ of Schramm’s question by deriving sharp estimates for the probability of half-plane arm events and planar j -arm events with $j > 1$, with power-law error bounds in most cases, which gives much more than what he asks for.

³Where the exponent is an integer, namely the half-plane 2- and 3-arm exponents and planar 5-arm exponent.

⁴Also independently obtained roughly at the same time by Binder, Chayes and Lei in [2].

⁵The exception is the planar one-arm case. See Remark 7.6 for more detail.

7.1.1 Main results

We start with necessary notation. Let \mathbb{T} denote the triangular lattice where each face is an equilateral triangle and \mathbb{T}^* denote the dual graph. We consider the critical Bernoulli percolation on \mathbb{T}^* in which each hexagon is colored red (=open) or blue (=closed) independently with equal probability. For $j \geq 1$, let $\mathcal{B}_j(r, R)$ denote the half-plane j -arm event that there exist j disjoint crossings with alternating colors of the semi-annulus $A^+(r, R)$ in the upper half-plane with inner and outer radii r and R respectively (see Section 7.3.3 for the precise definition of arm events and Section 7.3.1 for the discretization of domains). We will also consider a variant of half-plane arm events, denoted by $\mathcal{H}_j(r, R)$, which corresponds to j -alternating arms from the segment $[-r, r]$ to the semi-circle C_R^+ within the half-disk B_R^+ . Write $b_j(r, R) := \mathbb{P}[\mathcal{B}_j(r, R)]$ and define h_j similarly. In the seminal work [25], it is showed that for any $j \geq 1$, the sequence $b_j(r, R)$ has a power-law decay in R with exponent

$$\beta_j = j(j+1)/6, \quad \text{such that as } R \rightarrow \infty, \quad b_j(r, R) = R^{-\beta_j + o(1)}. \quad (7.1)$$

In [18] the following improvement of b_1 is obtained:

$$b_1(1, n) = e^{O(\sqrt{\log \log n})} n^{-1/3} = (\log n)^{O(1/\sqrt{\log \log n})} n^{-1/3}. \quad (7.2)$$

We are now ready to state our main results for half-plane arm probabilities, which give sharp asymptotics with power-law error bounds for both b_j and h_j .

Theorem 7.1. *For any $j \geq 1$, there exist constants $c(j, r) > 0$ and $C_b(j, r), C_h(j, r) \geq 0$ such that for any real $n > r$,*

$$b_j(r, n) = C_b n^{-\beta_j} (1 + O(n^{-c})); \quad h_j(r, n) = C_h n^{-\beta_j} (1 + O(n^{-c})). \quad (7.3)$$

Note that in the cases of $j = 1, 2, 3$, $C_b(j, 1), C_h(j, 1) > 0$, hence it makes sense to pick $r = 1$, and our theorem applies to classical half-plane arm events out of one hexagon. The same asymptotics also hold for any fixed “inner initial configuration”. In this case, the constants depend on j and the initial configuration.

We now turn to (polychromatic) planar j -arm events. Again, see Sections 7.3.1 and 7.3.3 for conventions and precise definitions. We start with the classical j -arm event $\mathcal{P}_j(r, R)$, which is defined as the event that there exist j disjoint crossings of the annulus $A(r, R)$ and not all of the same color (except in the case $j = 1$). We will also work with several variants of arm events, with

requirements on color sequences, location constraints on the outer ending points of arms and/or the connectedness of certain pairs of arms. In particular, for $j > 1$, we consider \mathcal{X}_j and \mathcal{A}_j , which are variants satisfying the requirement on color sequences, with and without location constraints respectively; see (7.14) and below for precise definition, as well as \mathcal{Y}_j and \mathcal{Z}_j , which are defined similarly but with the extra requirement on the connectedness; see (7.15) and below for precise definitions. In the cases $j = 2, 3, 4, 5, 6$, $\mathcal{A}_j(1, R)$ are the classical alternating j -arm events.

Write $p_j(r, R) = \mathbb{P}[\mathcal{P}_j(r, R)]$ and define a_j , x_j , y_j and z_j similarly. In [25], it is showed that for any $j \geq 2$, the sequence $p_j(r, R)$ has a power-law decay in R with critical exponent

$$\alpha_j = (j^2 - 1)/12, \quad \text{such that as } R \rightarrow \infty, \quad p_j(r, R) = R^{-\alpha_j + o(1)}. \quad (7.4)$$

We now state our main results for planar arm probabilities which give sharp asymptotics with power-law bounds for variants x_j and y_j and as a consequence sharp asymptotics without explicit error bounds for a_j and up-to-constants estimates for p_j .

Theorem 7.2. *For any $j \geq 2$, there exist constants $c(j, r)$ and $C_x(j, r), C_y(j, r) \geq 0$ such that for all real $n > r$,*

$$x_j(r, n) = C_x n^{-\alpha_j} (1 + O(n^{-c})); \quad y_j(r, n) = C_y n^{-\alpha_j} (1 + O(n^{-c})). \quad (7.5)$$

Note that when $j = 2, 3, 4, 5$, $C_x(j, 1), C_y(j, 1) > 0$, and also when $j = 6$, $C_x(6, 1) > 0$. Hence in these cases (note that special care is needed for $x_6(1, n)$; see Remark 7.43 for discussions) our results hold for arm events out of a single hexagon. As an immediate consequence of Theorem 7.2 and the up-to-constants asymptotic equivalence of different arm events (see Claim (2) of Lemma 7.11),

$$p_j(r, n) \asymp a_j(r, n) \asymp n^{-\alpha_j}. \quad (7.6)$$

Moreover, we are able to obtain sharper estimates (albeit without a power-law error bound) than (7.6) for the alternating arm probability $a_j(r, n)$. For simplicity, we only give them for arm events out of one hexagon for $j \leq 5$ (but similar asymptotics hold for arm events with other inner initial configurations, including the case of $a_6(1, n)$). See Remark 7.43 for more discussions.

Theorem 7.3. *For $j = 2, 3, 4, 5$, there exists some $C_a(j) > 0$ such that for all real $n > 1$,*

$$a_j(1, n) = C_a n^{-\alpha_j} (1 + o(1)). \quad (7.7)$$

A direct application of Theorem 7.3 for $j = 2, 4$ is that one can replace renormalizing factors in the scaling limit of pivotal and interface measures by precise powers of the mesh size in Theorems 4.3 and 5.5 of [8] and that in the natural parametrization in Theorem 1.4 of [10]. Another application is the improvement for the asymptotics of the correlation or characteristic length, which is also a central object in the study of near-critical and dynamical percolation. In particular, our asymptotics on planar 4-arm events imply that various versions of the correlation length for planar critical percolation on triangular lattice are up-to-constants equivalent to $|p-1/2|^{-4/3}$. For a more thorough account on the correlation length, see e.g., Section 7 of [19].

7.1.2 Comments

In this subsection, we briefly comment on the proof and discuss possible directions for generalization.

We start with the strategy of the proof. As Schramm has already points out immediately below his question in [22], the crux of the matter lies in “the passage between the discrete and continuous setting” and he poses a related question on obtaining “reasonable estimates for the speed of convergence of the discrete processes”.

This question is solved recently by Binder and Richards in [3], which constitutes the PhD thesis [20] of Richards; see also [4] for an extended abstract. More precisely, they verify that the framework developed in [26] for the power-law convergence rate of random discrete models towards SLE indeed works for percolation, the harmonic explorer and the FK-Ising model. In particular, they obtain a power-law convergence rate for the exploration process of planar critical percolation up to some stopping time. See Section 7.3.7 for detailed discussions of their results.

We now explain how we derive sharp asymptotics for arm events out of the power-law convergence. For better illustration we will discuss the classical half-plane arm-events (i.e., Theorem 7.1 for b_j) in more detail and only briefly point out necessary modifications in other cases.

At first glance, sharp asymptotics for arm events should follow naturally once we relate the discrete exploration path to SLE_6 . However, a few obstacles prevent us from applying the power-law convergence rate directly.

The first one is that such results only give information down to the mesoscopic level (in terms of the mesh size), while in this work we endeavor to reach the microscopic level. To overcome this difficulty, one resorts to coupling techniques developed for critical planar percolation, which allow us to “decouple” the microscopic “initial configurations” (if we figuratively regard arm events as

the consequence of outward explorations) with the macroscopic boundary conditions. Inspired by the arguments of Theorem 1.2 in [17] for the one-point function of 3D LERW, the ideas above can be crystallized in the following proportion estimates, from which Theorem 7.1 for b_j can be directly derived.

Denote $r_\star(j) = \min\{r \in \mathbb{N} : \star_j(r, n) > 0 \text{ for all sufficiently large } n\}$ for $\star \in \{b, h, p, a, x, y, z\}$. It is clear that when dealing with some arm probability $\star_j(r, n)$, it suffice to focus on the case $r \geq r_\star(j)$. Note that $r_b(j) \leq r_h(j)$ and $r_a(j) = r_x(j) \leq r_y(j) = r_z(j)$; in addition, $r_h(j) = 1$ for $j \leq 3$ and $r_y(j) = 1$ for $j \leq 5$.

Proposition 7.4. *Given $j \geq 1$, for any $r \geq r_b(j)$ and $m \in (1.1, 10)$,*

$$\frac{b_j(r, m^2 n)}{b_j(r, mn)} = \frac{b_j(r, mn)}{b_j(r, n)} \left(1 + O(n^{-c})\right), \quad (7.8)$$

where $O(n^{-c})$ is independent of the choice of m .

The second one is that the classical arm event \mathcal{B}_j is not the ideal choice to be described by the exploration path in a discretized domain. More precisely, working with \mathcal{B}_j directly involves dealing with a sequence of semi-annuli with shrinking inner radii. Although in principle one should be able to produce similar power-law convergence in these varying domains as a perturbation of the half-disk, a more sensible choice is to make use of couplings of percolation configurations conditioned on arm events (see the discussion on the proof of Proposition 7.4 below), and work with the variant \mathcal{H}_j instead; see Proposition 7.19 for more detail.

The third one is that the result of Binder and Richards provides convergence rate only for the exploration up to some stopping time, not for the whole path (although it is believed to hold for the latter as well). To overcome this difficulty, we enlarge the domain in which we compare percolation explorations and SLE_6 , so that the segment before the stopping time already suffices to provide comparison of arm probabilities; see Section 7.5.1 for more detail.

We now discuss the proof of Proposition 7.4 in more detail. The main ingredients are Propositions 7.28 and 7.29, in which we derive comparisons of conditioned arm probabilities, which allow us to compare arm probabilities between different boundary conditions and also to compare different types of arm events. These estimates are established through coupling results established in Propositions 7.25 and 7.27 for $j = 1$ and $j > 1$ respectively as well as the power-law convergence of the exploration proces in the form of Proposition 7.35.

A crucial ingredient for these couplings is the separation lemma, which, first developed by Kesten in [12], plays a key role in establishing quasi-multiplicativity properties of arm probabilities. We will discuss it (and several variants) in detail in Sections 7.3.5 and 7.3.6. It is also worth mentioning that although not directly needed in the proof, we also obtain a super-strong separation lemma for the half-plane setup, which confirms a conjecture by Garban, Pete and Schramm in [8] and is of independent interest.

We now turn to the plane case. Similarly, Theorem 7.2 for y_j can be derived from the following proportion estimates.

Proposition 7.5. *Given $j \geq 2$, for any $r \geq r_y(j)$ and $m \in (1.1, 10)$,*

$$\frac{y_j(r, m^2 n)}{y_j(r, mn)} = \frac{y_j(r, mn)}{y_j(r, n)} \left(1 + O(n^{-c})\right), \quad (7.9)$$

where $O(n^{-c})$ is independent of the choice of m .

This proposition follows from the same argument as that of Proposition 7.4 as the specific definition of \mathcal{Y}_j allows us to relate it to the exploration process. Note that in order to overcome the third obstacle as in the half-plane case, we will also enlarge the domain accordingly; see Section 7.5.3 for more detail. The claim for x_j follows from a coupling result that relates y_j and x_j ; see Proposition 7.34.

Theorem 7.3, whose proof is inspired by Proposition 4.9 of [8], is a corollary of Theorem 7.2, Proposition 7.33 (which relates the conditional arm probabilities for \mathcal{X}_j and \mathcal{A}_j), and Claim (4) of Lemma 7.11 (which gives the convergence of macroscopic arm probabilities without a speed).

Remark 7.6. We now briefly mention some open questions and discuss possible generalizations.

1) Planar 1-arm events are crucial objects to understand critical percolation clusters. In [16] it is shown that the 1-arm probabilities satisfy

$$p_1(r, R) = R^{-5/48 + o_R(1)} \quad \text{as } R \rightarrow \infty. \quad (7.10)$$

It is then a natural question to wonder if our asymptotics also hold in this case. In fact, the key difficulty in extending our arguments to the one-arm case lies in the fact that planar one-arm events cannot be described by a single chordal exploration path. Instead, they can be described either by the collection of all interface loops or by the so-called radial exploration process, which

correspond to CLE_6 or the radial SLE_6 respectively in the scaling limit. While we believe power-law convergence analogous to Theorem 4.1.11 in [20] should also hold for these objects, solid arguments leading to such results still seem rather out of reach for the moment.

2) Greater difficulties exist for the (planar) monochromatic arm-events. In [1], Beffara and Nolin show the existence of monochromatic arm exponents for critical planar percolation using quasi-multiplicativity arguments. Although couplings of conditional monochromatic arm events should exist in some form (if so, it must be proved in a way different than ours), the lack of an explicit characterization through the exploration path is the main difficulty that prevents us from obtaining any new results.

3) One may also wonder if our results hold for other lattices and/or types of percolation, in particular the critical bond percolation on \mathbb{Z}^2 . In this direction, we believe that all our coupling arguments (including the separation lemmas) will work with very little adaptation, while the biggest obstacle is definitely the passage from discrete to continuum. See [7] for some recent progress in this direction.

4) Another interesting direction is to obtain sharp asymptotics also for the FK-Ising arm events and the one-point function of the harmonic explorer (as well as the convergence in natural parametrization to SLE_4), for Binder and Richards also establish power-law convergence for these two models in [3]. We plan to investigate them in future works.

Finally, we explain how this article is organized. In Section 7.2, we settle the setup and introduce various key notions and preliminary results, including arm events, faces, separation lemmas, and the power-law convergence rate of the exploration paths. We give without proof various coupling results concerning arm events in Section 7.4 which are crucial to the main arguments and derive out of these couplings some relations on conditioned arm probabilities. Section 7.5 is dedicated to the proof of the main theorems in which we also include the proofs of Propositions 7.4 and 7.5. In Section 7.6, we unfold the proofs for various coupling results postponed from Section 7.4. A few technical proofs for preliminary results in Section 7.2 are given in the Appendices.

Acknowledgments: Hang Du, Yifan Gao and Xinyi Li thank National Key R&D Program of China (No. 2021YFA1002700 and No. 2020YFA0712900) and NSFC (No. 12071012) for support. Hang Du is also partially supported by the elite undergraduate training program of School of Mathematical Sciences at Peking University. Zijie Zhuang is partially supported by NSF grant DMS-1953848. Xinyi Li also thanks Daisuke Shiraishi for inspiring discussions.

7.2 Notation and preparatory results

This section is dedicated to setup and preparatory results. In Section 7.3 we introduce overall notation. We then fix the setup and recall some basic tools for percolation in Sections 7.3.1–7.3.2 resp. In Section 7.3.3, we give the definition for various arm events, state some existing results their asymptotics and give a “functional equation” type result on sequences that will be useful for obtaining sharp asymptotics. In Sections 7.3.5 and 7.3.6, we state and prove the separation lemmas in the half-plane and plane resp., which are crucial to the coupling argument in this work. Finally in Section 7.3.7, we recall the power-law convergence rate of the rescaled exploration path towards SLE_6 from [20] and prove a variant tailored for this work.

7.3 Notation and conventions

Let \mathbb{C} stand for the complex plane and $\mathbb{H} = \{x + iy : y > 0\}$ stand for the upper half-plane. For convenience of notation we regard them as \mathbb{R}^2 and a subset thereof and use both sets of notation interchangeably. Let $B(x, R) = \{z : |z - x| < R\}$ denote the ball of radius R around x and $C(x, R) = \partial B(x, R)$. For $0 < r < R$, we write $A(x, r, R) := \{z : r < |z - x| < R\}$ for the annulus of radii $r < R$ around x . We will omit “ x ” when $x = 0$, i.e., abbreviate $B(0, R), C(0, R), A(0, r, R)$ as $B_R, C_R, A(r, R)$. We will add $+$ in the superscript to indicate the quantities are defined in the upper half-plane \mathbb{H} . For example, if $A \subseteq \mathbb{C}$, write A^+ for the set $A \cap \mathbb{H}$. Then, $B^+(x, R), C^+(x, R), A^+(x, r, R), B_R^+, C_R^+, A^+(r, R)$ are the corresponding subsets of \mathbb{H} . Without further specification, all the sequences we consider are indexed by real numbers, not only integers.

We write $a_n = n^{\alpha+o(1)}$ if for all $\epsilon > 0$, $n^{\alpha-\epsilon} \leq a_n \leq n^{\alpha+\epsilon}$ for all n large enough. We write $a_n = O(n^\alpha)$ if there exists $C > 0$ such that $|a_n| \leq Cn^\alpha$ for all n , and write $a_n \asymp b_n$ if $a_n, b_n > 0$ for all n large enough, and a_n/b_n is bounded both from above and below. We always write j for the number of arms and η for the scale of lattice. Without further specification, all the arcs are written counterclockwise. We write $d(x, y)$ for the Euclid distance of $x, y \in \mathbb{R}^2$. Given a real number x , we write $\lfloor x \rfloor := \max_{z \in \mathbb{Z}} \{z \leq x\}$ for the integer part of x .

Finally, let us explain our convention concerning constants. Constants like $\varepsilon, \delta, c, c', C$ or C' may change from line to line while those with a subscript like c_1 are kept fixed throughout the paper. All constants are universal except those marked at the first occurrence.

7.3.1 Setting for percolation

Let \mathbb{T} be the triangular lattice embedded in the complex plane \mathbb{C} where each face is an equilateral triangle of side length 1. More precisely, the set of sites (vertices) in \mathbb{T} is given by $T := \mathbb{Z} + e^{i\pi/3}\mathbb{Z}$, and two sites are neighbors if they are at distance 1.

We consider the critical site percolation on \mathbb{T} , defined through declaring each site $v \in T$ as open or closed equally with probability $\frac{1}{2}$, independently of the other sites. More precisely, let $\{0, 1\}^T$ be the sample space of configurations $(\omega_v)_{v \in T}$, where $\omega_v = 1$ if v is open, and $\omega_v = 0$ if v is closed. Write \mathbb{P} for the product measure with parameter $\frac{1}{2}$ on T . For better illustration, a site will be colored red (resp. blue) if it is open (resp. closed). In studying scaling limits we will also be interested in percolation on the rescaled lattice. For $\eta > 0$, let $\eta\mathbb{T}$ be \mathbb{T} rescaled by η . With slight abuse of notation, we will also use \mathbb{P} to denote the product measure with parameter $\frac{1}{2}$ on ηT . However, we will fix $\eta = 1$ in the majority of this work and emphasize the mesh size only when⁶ we do rescale the lattice.

We also view the percolation on the triangular lattice \mathbb{T} as a random coloring of **hexagons** (i.e. faces) on \mathbb{T}^* , which is dual to \mathbb{T} such that each vertex (resp. edge) on \mathbb{T} corresponds to a hexagonal face (resp. an edge) on \mathbb{T}^* . We say two hexagons are **neighbors** (or adjacent) if they have a common edge. A **path** on \mathbb{T}^* is a sequence of neighboring hexagons and they are all distinct, and we call it a **loop** if the two ends of the path are also neighbors. We will also consider paths on \mathbb{T}^* as a graph, which we refer to as **b-paths** to distinguish them from paths of hexagons we have just defined⁷. We say two interfaces are **adjacent** if there is a hexagon on \mathbb{T}^* touched by both of them.

Let ∂_1 and ∂_2 be two disjoint parts of the boundary of a hexagonal domain D (in many cases two opposite sides of a topological quadrangle or inner and outer boundaries of a topological annulus). A **crossing** from ∂_1 to ∂_2 in D is a monochromatic path of hexagons in $\eta\mathbb{T}^*$ such that the two ends intersect ∂_1 and ∂_2 respectively, and all other hexagons are inside D . A crossing will also be called an **arm** when we consider arm events in what follows.

We now turn to the issues related to discretization of domains. By default, we use the conventions in [5] and refer readers thereto for more details. However, to avoid inconsistency for the discretization of (half)-disks, (semi-)annuli and (semi-)circles that are ubiquitous in this work, we

⁶Namely in Subsections 7.3.7, 7.5.1 and 7.5.3.

⁷In the language of this work, arms are paths while interfaces are b-paths.

give a special rule for the discretization of these objects; see the last paragraph of this subsection for details. We also remark that as will be discussed in Remark 7.24, our results are indeed quite stable against different choices of discretization. Hence our convention here is in fact merely an inessential technical matter.

We start with the default convention. A set of hexagons E is called connected or simply connected if so is the set obtained by embedding E in \mathbb{C} . If E is a simply connected set of hexagons, we write ΔE for the outer boundary of E , i.e., the set of hexagons that are not in E but adjacent to hexagons in E , and we write ∂E for the topological boundary of E by viewing E as a domain in \mathbb{C} . The bounded simply connected set E of hexagons is called a **Jordan set** if ΔE is a loop.

Given a Jordan domain D in \mathbb{C} , we write $D_{[\eta]}$ for the η -approximation of D , i.e., the largest Jordan set of $\eta\mathbb{T}^*$ that is contained in D . For $a \in \partial D$, we write a_η for the vertex of $\eta\mathbb{T}^*$ in $\partial D_{[\eta]}$ such that it is closest to a and it is on the common edge shared by two adjacent hexagons in $\Delta D_{[\eta]}$ (if there is a tie, choose one arbitrarily). When $\eta = 1$, we omit η or $[\eta]$ from the subscript unless we wish to emphasize that this is a discretized object.

We now turn to special discretization rules for points and intervals on the real axis, circles, disks and annuli centered at the origin. Note that if one applies the rules in the paragraph above to e.g., B_r^+ and $A(r, R)$ which are two touching objects in the continuum, the resulting hexagonal domains will have non-empty “cracks” in-between. Hence, we apply the following special rule for them.

We first designate, in an arbitrary way, the discrete circles⁸ C_R for each $R \geq 1$ as a loop of b-paths on \mathbb{T}^* that disconnects the origin from infinity which is at most $O(1)$ -away from C_R , such that C_1 is the boundary of the hexagon containing the origin and there do not exist $R_1 < R_2$ such that C_{R_1} crosses C_{R_2} . Then, for each $1 \leq r < R$ we regard B_R as the hexagonal domain encircled by C_R and $A(r, R)$ as $B_R \setminus B_r$ respectively. Similarly, we also designate for each $R \geq 1$ the discrete semi-circle C_R^+ as a b-path that disconnect 0_1 from infinity in $\mathbb{H}_{[1]}$ with the special rule that C_1^+ is the upper three edges of the hexagonal face in \mathbb{T}^* containing 0_1 . We then define half-disks B_R^+ and semi-annuli $A^+(r, R)$ in a similar fashion. With slight abuse of notation, we write $[-R, R]$ for the b-path that forms the part of the boundary of $\mathbb{H}_{[1]}$ encircled by C_R^+ and define other intervals

⁸With slight abuse of notation we still denote by C_R the discretized object; same for other objects in this paragraph.

accordingly. When working on the discretization with respect to the rescaled lattice $\eta\mathbb{T}$, we first counter-rescale to $\eta = 1$, apply the rules above and then rescale everything back to $\eta\mathbb{T}$.

7.3.2 Some classical tools

The aim of this subsection is to review some classical tools in percolation theory, namely, the Harris-FKG inequality, the BK-Reimer inequality and the Russo-Seymour-Welsh theory. These tools will be extensively used throughout this paper. We refer readers to [9] for more details.

Definition 7.1 (Monotone events). *There is a natural partial order on the set of configurations $\{0, 1\}^T$ given by $\omega_1 \leq \omega_2$ if and only if $\omega_1(v) \leq \omega_2(v)$ for all $v \in T$. An event A is called increasing if $1_A(\omega_1) \leq 1_A(\omega_2)$ whenever $\omega_1(v) \leq \omega_2(v)$, where 1_A is the indicator function of A . An event A is called decreasing if its complement A^c is increasing.*

Lemma 7.7 (Harris-FKG inequality). *If A and B are both increasing or decreasing events, then*

$$\mathbb{P}[A \cap B] \geq \mathbb{P}[A]\mathbb{P}[B].$$

Lemma 7.8 (BK-Reimer inequality). *Let $A \square B$ denotes the disjoint occurrence of the events A and B , meaning that $\omega \in A \square B$ if there exist disjoint sets of sites E and F (possibly depending on ω) with the property that one can verify that $\omega \in A$ (resp. $\omega \in B$) by looking at sites in E (resp. F) only. Then, we have*

$$\mathbb{P}[A \square B] \leq \mathbb{P}[A]\mathbb{P}[B].$$

Lemma 7.9 (Russo-Seymour-Welsh (RSW) Theorem). *For any topological quadrangle (i.e., a Jordan domain with four marked points on the boundary) D in \mathbb{C} with two opposite sides ∂_1 and ∂_2 , there exist constants $\eta_0 > 0$ and $c = c(D) > 0$ such that for all $\eta \leq \eta_0$, with probability at least c , there is an open crossing in $\eta\mathbb{T}^*$ from ∂_1 to ∂_2 in $D_{[\eta]}$ (which we call a quad crossing).*

The combination of the Harris-FKG inequality and RSW theorem allows us to “glue” arms of the same color to create paths that have macroscopic geometric restrictions with uniformly positive probability. E.g., a loop that disconnects the annulus $A(R, 2R)$ can be constructed through gluing of quad crossings. We will colloquially refer to these arguments as “RSW-FKG gluing” or “applications of the RSW theory”.

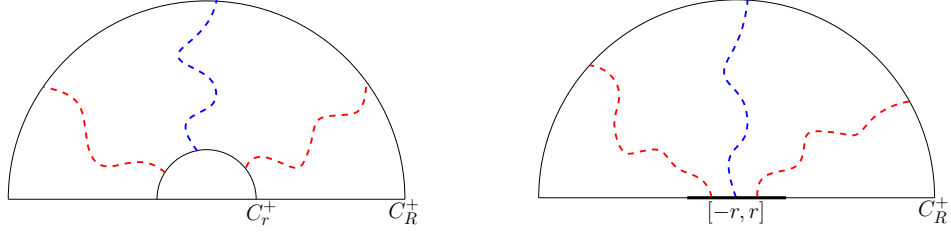


Figure 7.1: **Left:** The event $\mathcal{B}_3(r, R)$. **Right:** The event $\mathcal{H}_3(r, R)$. The semi-annulus $A^+(r, R)$ and the line segment $[-r, r]$ (in bold black) should be viewed as b-paths discretized according to the convention given at the end of Section 7.3.1. Dashed red and blue curves represent arms (composed of hexagons) with respective colors.

7.3.3 Arm events and asymptotics

In this subsection, we briefly review the definition of classical arm events in planar percolation and introduce variants that we will specifically use in this work. We then review existing asymptotics of the arm probabilities and finally lay out a “functional equation” lemma on sequences that is tailored for the derivation of sharp asymptotics in our work.

We first introduce two types of half-plane arm events. We fix $\eta = 1$ throughout these definitions. In the following, the symbols $B_R^+, C_R^+, A^+(r, R)$, $[-r, r]$, C_r and $A(r, R)$ refer to discrete objects introduced at the end of Section 7.3.1. For $j \geq 1$, let $\mathcal{B}_j(r, R)$ and $\mathcal{H}_j(r, R)$ stand for the half-plane j -arm events from scale r to scale R in the semi-annulus and half-disk respectively, defined as

$$\mathcal{B}_j(r, R) = \{\text{There are } j \text{ disjoint arms from } C_r^+ \text{ to } C_R^+ \text{ in } A^+(r, R) \text{ of alternating colors}\}; \quad (7.11)$$

$$\mathcal{H}_j(r, R) = \{\text{There are } j \text{ disjoint arms from } [-r, r] \text{ to } C_R^+ \text{ in } B_R^+ \text{ of alternating colors}\}, \quad (7.12)$$

where **alternating colors** means that the color pattern is red, blue, red, \dots , in counterclockwise order. See Figure 7.1 for an illustration.

Remark 7.10. We now briefly comment on the definitions above.

1) Among the two definitions above, $\mathcal{B}_j(r, R)$ is the classical half-plane arm event in the literature, however it is not the convenient setup for us to relate to the scaling limit of the exploration process. To overcome this difficulty, we introduce $\mathcal{H}_j(r, R)$, which perfectly solves this problem. See Section 7.5.1, in particular Lemma 7.37 for how the definition of $\mathcal{H}_j(r, R)$ comes into play.

2) As discussed in Remark 2 in [25], in the half-plane case it is not restrictive to study arm events of alternating colors thanks to a “color-switching” trick. In fact, the probability of $\mathcal{B}_j(r, R)$ (or $\mathcal{H}_j(r, R)$) remains the same for any color sequence.

We now introduce the (polychromatic) arm events in the plane. For $j \geq 2$, the classical j -arm event from scale r to R in the plane is given by

$$\mathcal{P}_j(r, R) = \{\text{There are } j \text{ disjoint arms from } C_r \text{ to } C_R \text{ in } A(r, R), \text{ not all of the same color}\}. \quad (7.13)$$

However, in this work we are going to mainly work with the following variants. We refer readers to Figure 7.2 for an illustration. Fix four points $a = (0, -1)$, $b = (-1/2, \sqrt{3}/2)$, $c = (1/2, \sqrt{3}/2)$ and $d = (0, 1)$ on C_1 . Let

$$l(j) = \lfloor j/4 \rfloor + \lfloor (j+1)/4 \rfloor + 1; \quad r(j) = \lfloor (j+2)/4 \rfloor + \lfloor (j+3)/4 \rfloor - 1.$$

(Note that $l(j) + r(j) = j$.) We call j disjoint arms from C_r to C_R in $A(r, R)$ are “with the prescribed pattern” if there are $l(j)$ of them from C_r to $R \cdot \widehat{ba}$ (the counterclockwise arc from $R \cdot b$ to $R \cdot a$) with color pattern red, blue, red \dots (counted clockwise from the point $R \cdot a$); and the rest $r(j)$ ones from C_r to $R \cdot \widehat{ac}$ with color pattern blue, red, blue \dots (counted counterclockwise from $R \cdot a$). Then, we define the following variants of planar j -arm events $\mathcal{X}_j, \mathcal{Y}_j$ as

$$\mathcal{X}_j(r, R) = \{\text{There exist } j \text{ disjoint arms from } C_r \text{ to } C_R \text{ in } A(r, R) \text{ with the prescribed pattern}\}, \quad (7.14)$$

$$\mathcal{Y}_j(r, R) = \{\text{There exist } j \text{ disjoint arms from } C_r \text{ to } C_R \text{ in } A(r, R) \text{ with the prescribed pattern}$$

and furthermore the left $(k+1)$ -th arm and the right k -th arm are connected

$$\text{in } B_R \text{ by a path of their (common) color for all } 1 \leq k \leq (l(j) - 1) \wedge r(j)\}. \quad (7.15)$$

We will also consider two other variants \mathcal{A}_j and \mathcal{Z}_j , which correspond to the events \mathcal{X}_j and \mathcal{Y}_j respectively but with no constraints on the endpoints on C_R . When j is even, \mathcal{A}_j is the arm event with alternating color sequences, referred to in the literature as the **alternating arm events**. Note that \mathcal{A}_j is different from \mathcal{P}_j except for $j = 2$, although they are comparable; See Claim (2) of Lemma 7.11.

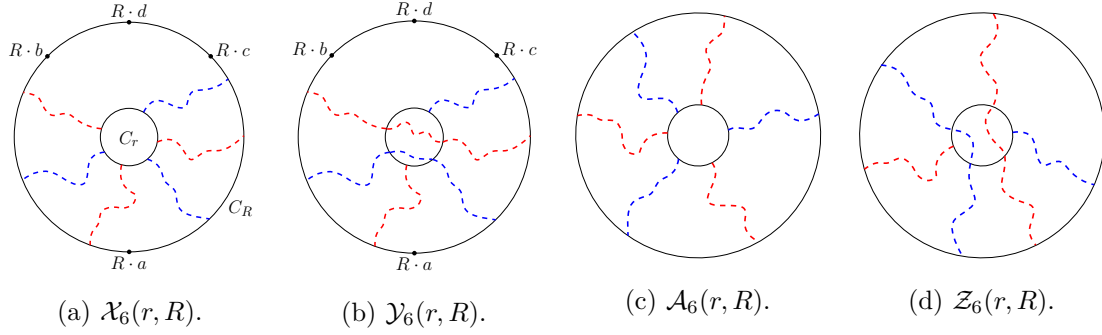


Figure 7.2: Illustration of the events $\mathcal{X}_j(r, R)$, $\mathcal{Y}_j(r, R)$, $\mathcal{A}_j(r, R)$ and $\mathcal{Z}_j(r, R)$ when $j = 6$. The annuli should be viewed as hexagonal domains on \mathbb{T}^* under the convention given at the end of Section 7.3.1. The dashed red and blue curves represent arms of the respective colors.

We now turn to arm probabilities. We write

$$b_j(r, R) = \mathbb{P}[\mathcal{B}_j(r, R)], \text{ and similarly } h_j, p_j, x_j, y_j \text{ and } a_j \text{ for arm probabilities.} \quad (7.16)$$

We now collect several quick facts and review some classical estimates on the asymptotics of arm probabilities. For simplicity, we will fix $r \geq r_h(j)$ and write $h_j(n)$ as a shorthand for $h_j(r, n)$ (and resp. for other arm probabilities) throughout this subsection.

We now collect several classical facts on the asymptotics of arm probabilities. In the following lemma, we show that in both half-plane and plane cases, the probabilities of variants arm events are up-to-constants equal to each other.

Lemma 7.11. *The half-plane arm probabilities are comparable for $j \geq 1$ and so is the case for planar arm events and $j \geq 2$. Namely*

- (1). $h_j(n) \asymp b_j(n)$;
- (2). $p_j(n) \asymp x_j(n) \asymp a_j(n)$ and for $r > r_0(j)$ such that $y_j(n) > 0$ for all large n , we further have $a_j(n) \asymp y_j(n)$.

Moreover, with the arm exponents defined in (7.1) and (7.4) resp.,

(3). $b_j(R_1, R_2) = (R_1/R_2)^{-\beta_j+o(1)}$ as $R_1 \geq r_b(j)$ and $R_1/R_2 \rightarrow 0$. Also for $j \geq 2$, $p_j(R_1, R_2) = (R_1/R_2)^{-\beta_j+o(1)}$ as $R_1 \geq r_p(j)$ and $R_1/R_2 \rightarrow 0$. The same holds for h_j, a_j, x_j, y_j .

Finally, the macroscopic arm probabilities have the following asymptotics:

(4). There exist $f_j, g_j : (0, 1) \rightarrow \mathbb{R}^+$ such that for $0 < \epsilon < 1$,

$$\lim_{n \rightarrow \infty} x_j(\epsilon n, n) = f_j(\epsilon) \quad \text{and} \quad \lim_{n \rightarrow \infty} a_j(\epsilon n, n) = g_j(\epsilon).$$

The proofs for Claims (1) and (2) are classical applications of separation lemmas (which we will introduce shortly in this section) and will be postponed to Appendix 7.7.1. Claim (3) is the main result in [25], while Claim (4) for a_j is Lemma 2.9 of [8] while the case for x_j follows from essentially same arguments.

Remark 7.12. The power of (3) lies in the fact that R_1 can also go to infinity with R_2 as long as R_1/R_2 goes to zero. In the context, we will use two particular forms frequently. One is that for fixed $\alpha \in (0, 1)$

$$h_j(n^\alpha, n) \asymp b_j(n^\alpha, n) = n^{-\beta_j(1-\alpha)+o(1)} \quad \text{and} \quad x_j(n^\alpha, n) \asymp y_j(n^\alpha, n) \asymp p_j(n^\alpha, n) = n^{-\alpha_j(1-\alpha)+o(1)}, \quad (7.17)$$

The other is that there exists $c > 0$ such that when $R_1/R_2 \rightarrow \infty$

$$b_3(R_1, R_2) \leq (R_1/R_2)^{1+c} \quad \text{and} \quad p_6(R_1, R_2) \leq (R_1/R_2)^{2+c},$$

because $\beta_3 = 2 > 1$ and $\alpha_6 = 35/12 > 2$.

In the following lemma, we show that the arm probabilities are “almost” monotone. We also postpone its proof to Appendix 7.7.1.

Lemma 7.13. For any $j \geq 1$

$$\lim_{\epsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{n \leq s \leq t \leq (1+\epsilon)n} \frac{h_j(t)}{h_j(s)} \geq 1.$$

The same inequalities also hold for b_j and x_j, y_j, p_j and a_j (for $j \geq 2$).

The following lemma is a “functional equation” type result on sequences and it shows that we can obtain Theorem 7.1 from Proposition 7.4 and Theorem 7.2 from Proposition 7.5 together with the preliminary estimates Lemma 7.11 and 7.13 (which imply that the sequence of various types of arm probabilities satisfy Assumption (2) in the following lemma). Its proof can be found in Appendix 7.7.2.

Lemma 7.14. *Consider a set of positive real numbers $\{a_n : n \in \mathbb{R} \cap [1, \infty)\}$ such that:*

- (1). $\frac{a_{m^2n}}{a_{mn}} = \frac{a_{mn}}{a_n} \left(1 + O(n^{-c})\right)$ for all $m \in (1.1, 10)$, where the constants in $O(n^{-c})$ is independent of m in the given range.
- (2). $\lim_{\epsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \inf_{n \leq s \leq t \leq (1+\epsilon)n} \frac{a_t}{a_s} \geq 1$.

Then, there exist $0 < C < \infty$ and $-\infty < \alpha < \infty$ such that $a_n = Cn^\alpha \left(1 + O(n^{-c})\right)$.

7.3.4 Faces

When working on couplings of arm events, a crucial concept we will need throughout the arguments is **faces**, which forms a special example of stopping sets. The notion of stopping set is in some sense a two-dimensional version of the stopping time.

We begin with faces in the half-plane. Recall the conventions given at the end of Section 7.3.1 for discretization. We call a set of paths $\Theta = \{\theta_1, \dots, \theta_j\}$ a configuration of **outer faces** around C_r^+ if the following conditions are satisfied:

- For $1 \leq i \leq j$, if i is odd (resp. even), then θ_i is a red (resp. blue) path from h_i to h^i such that $\theta_i \subset \mathbb{H} \setminus B_r^+$.
- For $1 \leq i \leq j$, the end-hexagons h_i 's and h^i 's satisfy the condition that h_1, h^j touch $\partial\mathbb{H}_\eta$, and $h^1, h_2, \dots, h^{j-1}, h_j$ touch C_r^+ . Furthermore, $(h_1, h^1, \dots, h_j, h^j)$ are listed in counterclockwise order, and h^i is adjacent to h_{i+1} for $1 \leq i \leq j-1$.

See Figure 7.3 for an illustration. The paths $\theta_1, \dots, \theta_j$ are called the **outer faces** of Θ . The set of points $x^i := h^i \cap h_{i+1} \cap C_r^+$ for all $1 \leq i \leq j-1$ are called the **endpoints** of Θ . The concepts of configurations of **inner faces** around C_r are defined similarly except that we require instead that $\theta_i \subset B_r^+$ for all i . In both cases, each path will be called a face. We will also use Θ to denote the union of all hexagons in $\theta_i, 1 \leq i \leq j$ with a slight abuse of notation. For Θ a configurations of outer faces, we write \mathcal{D}_Θ and \mathcal{V}_Θ for the connected components of $\mathbb{H}_\eta \setminus \Theta$ whose boundary contains ∞ and 0 respectively⁹. When Θ represents a configuration of inner faces, we exchange the role of ∞ and 0 in the definition accordingly. See Figure 7.3 for an illustration.

⁹In these definitions, the initials D and V stand for “discovered” and “vacant”, signifying the status of these regions when we explore inwards; same for inner faces. See also the definition for stopping sets.

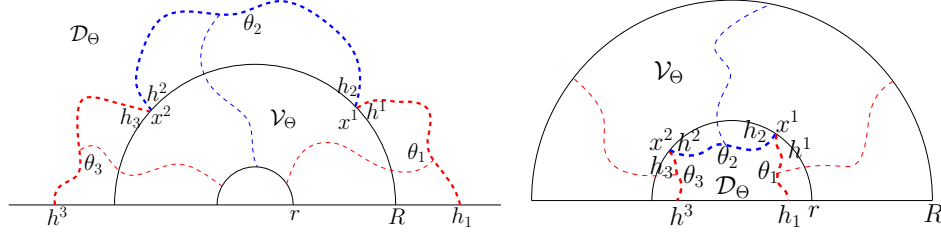


Figure 7.3: **Left:** A configuration of outer faces around C_R^+ . **Right:** A configuration of inner faces around C_r^+ . In both cases, $j = 3$ and the configuration of faces $\Theta = \{\theta_1, \theta_2, \theta_3\}$ is depicted in heavy dashed curves with corresponding colors. We have indicated the relative position of \mathcal{D}_Θ and \mathcal{V}_Θ . The events $\mathcal{B}_j^\Theta(r, R)$ occur when the arms indicated by normal dashed curves with colors exist. Moreover, h_i and h^i are the end-hexagons of θ_i , and x^1 and x^2 are the endpoints Θ .

Suppose $r < R$. If Θ is a configuration of outer faces around C_R^+ , then we write $\mathcal{B}_j^\Theta(\mathbf{r}, \mathbf{R})$ (resp. $\mathcal{H}_j^\Theta(\mathbf{r}, \mathbf{R})$) for the event that each outer face is connected to C_r^+ (resp. $[-r, r]$) by an arm in \mathcal{V}_Θ of the same color. The same definition applies to a configuration of inner faces Θ around C_r^+ , with each face connected to C_R^+ by an arm also in \mathcal{V}_Θ . When Θ is specified, there would be no confusion in whether $\mathcal{B}_j^\Theta(\mathbf{r}, \mathbf{R})$ refers to a configuration of inner or outer faces.

In a similar fashion, in the plane, we define the configuration of outer or inner faces around a circle as a circular chain of j (an even integer) paths of alternating colors. More precisely, we call the set of paths $\Theta = \{\theta_1, \dots, \theta_j\}$ a configuration of outer (resp. inner) faces around C_r if

- $\theta_1, \dots, \theta_j$ have alternating colors and they are contained in $\mathbb{C} \setminus B_r$ (resp. B_r).
- The end-hexagons of these j paths, $(h_1, h^1, \dots, h_j, h^j)$, touching C_r , are listed in counter-clockwise order (still, h_i and h^i are the two ends of θ_i). Moreover, h^i is adjacent to h_{i+1} for all $1 \leq i \leq j$ (where we set $h_{j+1} := h_1$).

See Figure 7.4 for an illustration. Again, the paths θ_i 's are called the outer (resp. inner) faces of Θ , and we can define the end-points x^1, \dots, x^j analogously (by setting $x^i := h^i \cap h_{i+1} \cap C_r$). Similar to the half-plane case, for any configuration of outer faces Θ , let \mathcal{D}_Θ and \mathcal{V}_Θ denote the connected components of $\mathbb{C} \setminus \Theta$ which contains ∞ and 0 respectively. If Θ stand for a configuration of inner faces instead, we then exchange the role of 0 and ∞ in the definition accordingly. For $0 < r < R$,

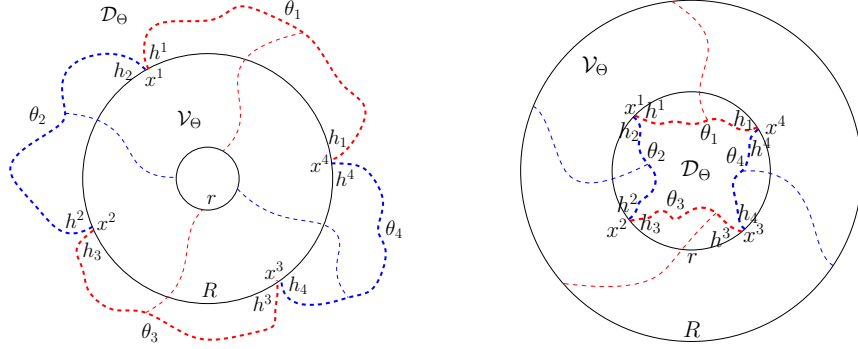


Figure 7.4: **Left:** A configuration of outer faces around C_R . **Right:** A configuration of inner faces around C_r . In both cases, $j = 4$ and the configuration of faces $\Theta = \{\theta_1, \dots, \theta_4\}$ is depicted in heavy dashed curves with corresponding colors. We have indicated the relative position of \mathcal{D}_Θ and \mathcal{V}_Θ . The events $\mathcal{A}_j^\Theta(r, R)$ occur when the arms indicated by normal dashed curves with colors exist. Moreover, h_i and h^i are the end-hexagons of θ_i , and x^1, \dots, x^4 are the endpoints of Θ .

if Θ is a configuration of outer faces around C_R , we write $\mathcal{A}_j^\Theta(\mathbf{r}, \mathbf{R})$ for the event that each outer face is connected to C_r by an arm of the corresponding color in \mathcal{D}_Θ . The corresponding events for inner faces are defined in a similar way.

By a **stopping set**, we mean a random configuration of inner or outer faces Θ which can be determined by an “one-sided” exploration process. That is to say, there exists an exploration process such that when the exploration stops, we are able to determine Θ from the configuration in \mathcal{D}_Θ only, and all the hexagons in \mathcal{V}_Θ are left unexplored.

7.3.5 Separation lemmas in the half-plane

In this subsection, we give different versions of the separation lemma in the half-plane. Their counterparts in the plane will be given in the next subsection separately.

The separation lemma for planar percolation first appears in Kestens work [12]. This is the key tool that allows one to deal with conditional arm probabilities and obtain quasi-multiplicativity for arm events. It is worth mentioning that its counterpart in the setup of Brownian motion established by Lawler in [15] also has multiple implications in the study of special points of Brownian motion.

We will state the separation lemma in terms of interfaces. By an **interface** we mean a b-path that separates clusters of different colors in some domain.

We now introduce the notion of “quality” to measure how separate the interfaces are. Suppose $u < v$. Let Γ be a set of interfaces crossing $A^+(u, v)$ from C_u^+ to C_v^+ . Suppose that Γ contains $j \geq 1$ interfaces and let $\{x^1, \dots, x^j\}$ be the collection of endpoints of these interfaces on C_v^+ listed in counterclockwise order. Then Γ has an **exterior quality** defined as

$$Q_{\text{ex}}(\Gamma) := \frac{1}{v} d(v, x^1) \wedge d(x^1, x^2) \wedge \dots \wedge d(x^j, -v). \quad (7.18)$$

In an analogous way, we define the **interior quality** for interfaces and denote it by Q_{in} . We can also define the quality¹⁰ Q for a configuration of faces Θ in a similar fashion. We call Θ to be **well-separated** if $Q(\Theta) > j^{-1}$.

The separation lemma roughly shows that conditioned on the interfaces reach a long distance without intersecting each other, then their endpoints will separate at some macroscopic distance with a universal positive probability. We now give the first version of separation lemma, which involves no initial configuration of faces, and is rather standard. This version can be viewed as the half-plane counterpart of [23, Lemma A.4] and its proof follows from the same line, so we omit it.

Lemma 7.15 (The Standard Separation lemma). *For any $j \geq 2$, there exists $c(j) > 0$ such that for any $100j \leq 2r \leq R$, the following hold.*

- Let Γ be the set of interfaces crossing $A^+(r, R)$ connecting C_r^+ and C_R^+ , then,

$$\mathbb{P}[Q_{\text{in}}(\Gamma) \wedge Q_{\text{ex}}(\Gamma) > j^{-1} \mid \mathcal{B}_j(r, R)] > c. \quad (7.19)$$

- Let Γ be the set of interfaces in B_R^+ connecting $[-r, r]$ and C_R^+ , then

$$\mathbb{P}[Q_{\text{ex}}(\Gamma) > j^{-1} \mid \mathcal{H}_j(r, R)] > c. \quad (7.20)$$

The following strong separation lemma is the half-plane version of [8, Proposition A.1]. In fact, Proposition 7.16 is called “strong” because it holds for any given initial configuration of faces which might be of very bad quality, not because it implies Lemma 7.15 (in fact it does NOT – note that in this strong version we do require the initial configuration to be j faces, which does not appear in the conditioning in Lemma 7.15). For the sake of completeness, we will prove it in Appendix 7.7.3.

¹⁰There is no need to add subscripts since in this case the type of quality can be read from the type of faces.

Proposition 7.16 (Strong separation lemma). *For any $j \geq 2$, there exists $c(j) > 0$ such that for any $100j\eta \leq 2r \leq R$, the following hold.*

- (Outward) *For any configuration of inner faces $\Theta = \{\theta_1, \dots, \theta_j\}$ with endpoints $\{x^1, \dots, x^{j-1}\}$ around C_r^+ , let Γ be the $(j-1)$ -tuple of interfaces in \mathcal{V}_Θ which start from x^1, \dots, x^{j-1} respectively until they reach C_R^+ . Then*

$$\mathbb{P}[Q_{\text{ex}}(\Gamma) > j^{-1} \mid \mathcal{B}_j^\Theta(r, R)] > c. \quad (7.21)$$

- (Inward) *In a similar fashion, the same claim also holds for any configuration of outer faces Θ with C_r^+ , C_R^+ and Q_{ex} replaced by C_R^+ , C_r^+ and Q_{in} respectively.*

In this work, we will also need a slightly strengthened version of the strong separation lemma stated below. Compared with Proposition 7.16, it lifts the restriction of requiring exactly j faces as the initial configuration, although it still requires an a priori upper bound on the number of faces. For a color configuration ω_0 on \mathbb{T}^* (or only on a subset D of \mathbb{T}^*), we denote by $\{\omega_D = \omega_0\}$ the event that the color configuration inside D coincides with ω_0 .

Proposition 7.17 (Slightly stronger separation lemma). *For any $K \geq j \geq 2$, there exists $c = c(j, K) > 0$ such that for any $10j \cdot 2^K \leq 2^K r \leq 2^K R_1 \leq R_2 \leq R$, letting Γ be the set of interfaces crossing $A^+(R_1, R_2)$, the following holds:*

- (Outward) *For any configuration of inner faces Θ around $C_{R_1}^+$ with no more than K faces, and any color configuration ω_0 that coincides with Θ and satisfies $\mathbb{P}[\mathcal{B}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$,*

$$\mathbb{P}[|\Gamma| = j - 1 \text{ and } Q_{\text{ex}}(\Gamma) > j^{-1} \mid \mathcal{B}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] > c. \quad (7.22)$$

- (Inward) *In a similar fashion, for any configuration of outer faces Θ around $C_{R_2}^+$ with no more than K faces, and any suitable color configuration ω_0 ,*

$$\text{The same inequality holds with } Q_{\text{ex}} \text{ replaced by } Q_{\text{in}}. \quad (7.23)$$

The claims (7.22) and (7.23) above also hold if $\mathcal{B}_j(r, R)$ is replaced by $\mathcal{H}_j(r, R)$.

Proof. We focus on (7.22) as all other cases are almost the same. Given any configuration of inner faces Θ around $C_{R_1}^+$ with k ($j \leq k \leq K$) faces, and any configuration ω_0 that coincides with Θ and satisfies $\mathbb{P}[\mathcal{B}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$, denote

$$\mathcal{B} = \mathcal{B}_j(r, R), \quad \mathcal{Q} = \{\omega_{\mathcal{D}_\Theta} = \omega_0\} \text{ and } \mathcal{U} = \{|\Gamma| = j - 1 \text{ and } Q_{\text{ex}}(\Gamma) > j^{-1}\}$$

for conciseness. Our aim is to show $\mathbb{P}[\mathcal{U} \mid \mathcal{B} \cap \mathcal{Q}] > c(j, K)$ for some constant $c(j, K) > 0$.

We first deal with the case $R_2 = R$ and we proceed by induction on K . The case $K = j$ reduces to (7.21) since when Θ has exactly j faces, $\mathbb{P}[\cdot \mid \mathcal{B} \cap \mathcal{Q}]$ is identical to $\mathbb{P}[\cdot \mid \mathcal{B}_j^\Theta(r, R)]$ in $\mathcal{V}(\Theta)$. Now, for some $K > j$, assume the result (when $R_2 = R$) is true for $(K - 1)$ and Θ has exactly K faces. Let Γ_1 be the set of interfaces starting from the $(K - 1)$ end-points of Θ , truncated at their first visit of $\mathbb{R} \cup C_{2R_1}^+$. Consider the event

$$\mathcal{B}_K^\Theta(R_1, 2R_1) = \{\text{all of these } (K - 1) \text{ interfaces reach } C_{2R_1}^+\}.$$

Assume $\mathcal{B}_K^\Theta(R_1, 2R_1)$ fails, then once we explore the $(K - 1)$ interfaces starting from the $(K - 1)$ endpoints and stop exploring immediately upon reaching $\mathbb{R} \cup C_{2R_1}^+$, we will see that some of them merge together or hit the real line before reaching $C_{2R_1}^+$. Thus the exploring process induces a configuration of inner faces Θ' around $C_{2R_1}^+$ with no more than $(K - 1)$ faces. Hence, the induction hypothesis gives

$$\mathbb{P}\left[\mathcal{U} \mid \mathcal{B}, \mathcal{Q}, (\mathcal{B}_K^\Theta(R_1, 2R_1))^c\right] > c(j, K - 1). \quad (7.24)$$

On the other hand, assume $\mathcal{B}_K^\Theta(R_1, 2R_1)$ happens. Clearly

$$\mathcal{B} \cap \mathcal{B}_K^\Theta(R_1, 2R_1) \subset \mathcal{B}_K^\Theta(R_1, 2R_1) \cap \mathcal{B}_j(4R_1, R),$$

and also $\mathcal{B}_j(4R_1, R)$, $\mathcal{B}_K^\Theta(R_1, 2R_1)$ and \mathcal{Q} are mutually independent, so

$$\mathbb{P}\left[\mathcal{B}, \mathcal{B}_K^\Theta(R_1, 2R_1) \mid \mathcal{Q}\right] \leq \mathbb{P}\left[\mathcal{B}_K^\Theta(R_1, 2R_1)\right] \mathbb{P}\left[\mathcal{B}_j(4R_1, R)\right]. \quad (7.25)$$

Furthermore, let Γ_2 be the set of interfaces crossing $A^+(4R_1, R)$, by FKG-RSW gluing,

$$\mathbb{P}\left[\mathcal{U}, \mathcal{B}, \mathcal{B}_K^\Theta(R_1, 2R_1) \mid \mathcal{Q}\right] \geq c_1 \mathbb{P}\left[\mathcal{B}_K^\Theta(R_1, 2R_1), Q_{\text{ex}}(\Gamma_1) > K^{-1}\right] \mathbb{P}\left[\mathcal{B}_j(4R_1, R), Q_{\text{in}}(\Gamma_2) > j^{-1}\right] \quad (7.26)$$

for some $c_1 = c_1(j, K) > 0$ (here the condition $\mathbb{P}[\mathcal{B}_j^\Theta \mid \mathcal{Q}] > 0$ assures the existence of feasible connecting pattern). By (7.19) and (7.21) (taking j therein as K), we get

$$\begin{aligned} \mathbb{P}\left[\mathcal{B}_K^\Theta(R_1, 2R_1), Q_{\text{ex}}(\Gamma_1) > K^{-1}\right] &> c_2 \mathbb{P}\left[\mathcal{B}_K^\Theta(R_1, 2R_1)\right] \text{ and} \\ \mathbb{P}\left[\mathcal{B}_j(4R_1, R), Q_{\text{in}}(\Gamma_2) > j^{-1}\right] &> c_3 \mathbb{P}\left[\mathcal{B}_j(4R_1, R)\right] \end{aligned}$$

for some positive constants $c_2(K), c_3(j)$. This combined with (7.25) and (7.26) yields

$$\mathbb{P}\left[\mathcal{U} \mid \mathcal{B}, \mathcal{Q}, \mathcal{B}_K^\Theta(R_1, 2R_1)\right] = \frac{\mathbb{P}\left[\mathcal{V}, \mathcal{B}, \mathcal{B}_K^\Theta(R_1, 2R_1) \mid \mathcal{Q}\right]}{\mathbb{P}\left[\mathcal{B}, \mathcal{B}_K^\Theta(R_1, 2R_1) \mid \mathcal{Q}\right]} > c_1 c_2 c_3 = c_4(j, K). \quad (7.27)$$

Now pick $c(j, K) = c(j, K-1) \wedge c_4(j, K)$, then $\mathbb{P}[\mathcal{U} \mid \mathcal{B}, \mathcal{Q}] > c(j, K)$ follows from (7.24), (7.27) and the total probability formula. This proves (7.22) for the case $R_2 = R$.

For the general case of (7.22), by the quasi-multiplicativity (see [23, (3.14)]) and what we just proved, we have for some constants $c_5, c_6, c_7 > 0$ depending only on j, K ,

$$\mathbb{P}[\mathcal{U}, \mathcal{B} \mid \mathcal{Q}] > c_5 \mathbb{P}[\mathcal{U}, \mathcal{B}_j(r, R_2) \mid \mathcal{Q}] \mathbb{P}[\mathcal{B}_j(R_2, R)] > c_6 \mathbb{P}[\mathcal{B}_j(r, R_2) \mid \mathcal{Q}] \mathbb{P}[\mathcal{B}_j(R_2, R)] > c_7 \mathbb{P}[\mathcal{B} \mid \mathcal{Q}].$$

This shows $\mathbb{P}[\mathcal{U} \mid \mathcal{B}, \mathcal{Q}] > c_7(j, K)$, as required. \square

Remark 7.18. The constant 2 in the condition $2r \leq R$ in the statement of Lemma 7.15 and Proposition 7.16 is not essential. Indeed, similar results still hold for any constant greater than 1 when r is large enough, and the proof is just a verbal change of the proof given in Appendix 7.7.3. Then it is clear from the above proof that Proposition 7.17 still holds for large r when the constant 2^K is replaced by any constant greater than 1.

Finally, in the half-plane, we are also able to obtain a much stronger version of the separation lemma, which first appears in the name of “super strong separation lemma” as Conjecture A.4 in [8] in a slightly different form. This “ultimate” version no longer requires any a priori upper bound on the number of faces and allows us to explore the configuration annulus by annulus when we construct couplings in Section 7.4, which is simpler and more natural than the exploration of faces.

Proposition 7.19 (Super strong separation lemma). *For any $j \geq 2$, there exist $M, c > 0$ depending only on j , such that for any $10M^2j \leq M^2r \leq R$ and $Mr \leq u \leq R/M$, if we let Γ be the set of interfaces crossing $A^+(r, R)$, then the following holds.*

- (Outward) *For any percolation configuration ω_0 which satisfies $\mathbb{P}[\mathcal{B}_j(r, R) \mid \omega_{B_u^+} = \omega_0] > 0$,*

$$\mathbb{P} \left[|\Gamma| = j - 1 \text{ and } Q_{ex}(\Gamma) > j^{-1} \mid \mathcal{B}_j(r, R), \omega_{B_u^+} = \omega_0 \right] > c.$$

- (Inward) *In a similar fashion, for any suitable color configuration ω_0 , the same claim follows with $\omega_{B_u^+}$ and Q_{ex} replaced by $\mathbb{H} \setminus B_u^+$ and Q_{in} respectively.*

In addition, similar results also hold if $\mathcal{B}_j(r, R)$ is replaced by $\mathcal{H}_j(r, R)$.

With this powerful tool in hand, we are able to derive the coupling results and estimates in the half-plane in a relatively straightforward (and natural) way. Although we believe that the super

strong separation lemma also holds for the plane case, unfortunately we could not find a proof due to the lack of a natural ordering for interfaces in the plane. In order to state and prove all results in parallel, even in the half-plane case we choose to use only the weaker version Proposition 7.17, which can be extended to the plane case (see Proposition 7.20) without difficulty. In what follows, we provide a sketch of proof for the half-plane super strong separation lemma.

Sketch of Proof for Proposition 7.19. We only discuss the outward case. The crux is to show the following fact: there is a constant $c'(j) > 0$, such that for any $Mr \leq u \leq R/M$ and any percolation configuration ω_0 satisfying $\mathbb{P}[\mathcal{B}_j(r, R) \mid \omega_{B_u^+} = \omega_0] > 0$ if we let Γ_u be the set of interfaces crossing $A^+(u, Mu/2)$, then $\mathbb{P}[|\Gamma_u| = j - 1 \mid \mathcal{B}_j(r, R), \omega_{B_u^+} = \omega_0] > c'$. Once this is true, Proposition 7.19 follows by exploring all interfaces crossing $A^+(u, Mu/2)$ then applying the strong separation lemma (Proposition 7.16) in $A^+(Mu/2, Mu)$.

To prove the fact, we note that the event $\{|\Gamma_u| > j - 1\}$ implies that there is a “pocket” down to scale u induced from some interface (see \mathcal{P}_1 in (a) of Figure 7.5). Basically, the existence of such an interface pocket together with some FKG-RSW gluing technique leads to the disjoint occurrence of $\mathcal{B}_j(r, R)$ and a rare one-arm event crossing a scale M_0 (see (a) and (b) of Figure 7.5). If this is the case, then applying BK-Reimer’s inequality yields the desired result. However, the interfaces might touch one another, imposing difficulty to find an extra arm and apply BK-Reimer’s inequality (see (c) of Figure 7.5). To tackle this issue, we need an iterative argument to deal with the nesting cases. Thanks to the simply connecting property of the zero-removed half-plane, we have a natural ordering for all these interfaces so the nesting process is indeed finite in some sense. With this observation, the desired fact can be proved by a union bound if we set M_0 large enough and choose $M = 2M_0^j$ such that the iterative arguments end in $(j - 1)$ steps. This concludes the proof. \square

7.3.6 Separation lemmas in the plane

This subsection is devoted to separation lemmas in the plane. As the arguments are similar to the half-plane case, we only give the statements in Proposition 7.20 and omit the proofs.

To begin with, we give the definition of the quality of interfaces, just as in the half-plane case. For $u < v$, let Γ be a set of percolation interfaces that start from C_u and end at the first visit of C_v . Suppose that Γ contains j interfaces with endpoints x^1, \dots, x^j on C_v , then we define the **exterior**

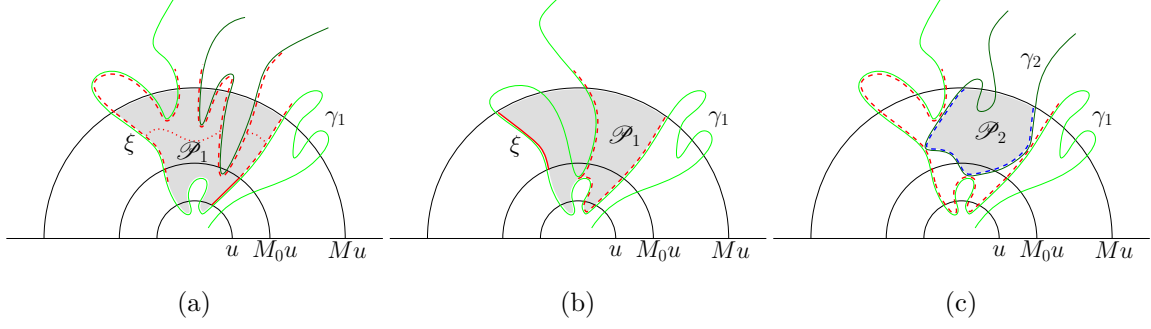


Figure 7.5: Three cases for the proof of Proposition 7.19. **(a)**: No touchings for \mathcal{P}_1 . Here and in case (b), γ_1 is the first (in counterclockwise order) global interface that crosses $A^+(u, Mu)$ more than once, depicted in green, while ξ is the leftmost crossing of γ_1 from u to Mu . The pocket \mathcal{P}_1 induced from γ_1 down to scale u is given by the gray region, which is enclosed by ξ and the part of γ_1 proceeding ξ from the last visit on C_{Mu}^+ . The figure illustrates the case that neither the part of γ_1 following ξ nor other (global) interfaces **touch** \mathcal{P}_1 (i.e., adjacent to \mathcal{P}_1 inside $A^+(u, M_0u)$). By RSW theory with positive probability there are red crossings (in dotted) connecting left and right sides of \mathcal{P}_1 , on which there is an extra red arm (in solid) that does not use the hexagons ensuring $\mathcal{B}_j(r, R)$. **(b)**: The interface γ_1 touches \mathcal{P}_1 . In this case, the red arm (in solid) neighboring ξ is an extra arm which plays the same role as that in case (a). **(c)**: Another interface touches \mathcal{P}_1 . We use γ_2 to denote this new interface which is depicted in dark green. Then γ_2 induces a pocket \mathcal{P}_2 inside \mathcal{P}_1 down to scale M_0u , depicted in gray (erasing \mathcal{P}_1). In the pocket \mathcal{P}_2 , we can iterate the arguments as before by considering existence (or non-existence thereof) of touchings of \mathcal{P}_2 inside $A^+(M_0u, M_0^2u)$.

quality of Γ as

$$Q_{\text{ex}}(\Gamma) := \frac{1}{v} d(x^1, x^2) \wedge d(x^2, x^3) \wedge \cdots \wedge d(x^j, x^1). \quad (7.28)$$

Again, we define the **interior quality** for interfaces in the same fashion and denote it by Q_{in} as well as the quality for a configuration of faces Θ and denote it by $Q(\Theta)$. If Θ satisfies $Q(\Theta) > j^{-1}$, we call it **well-separated**. Recall the notation $\mathcal{V}(\Theta)$ and $\mathcal{D}(\Theta)$ from Section 7.3.4.

There are counterparts of Lemma 7.15 and Proposition 7.16 in the plane case, but things become subtler since the number of interfaces crossing an annulus must be even. Lemma A.4 and Proposition A.1 in [8] only consider the case when j is even, but the same argument with little change yields similar results for odd j . More precisely, we set $J = 2\lfloor j/2 \rfloor$ for any $j \geq 2$, then there exists $c(j) > 0$ such that for any $10j \leq 2r \leq R$, and any configuration of outer faces $\Theta = \{\theta_1, \dots, \theta_J\}$ around C_R , let Γ be the set of interfaces crossing $A(r, R)$, then

$$\mathbb{P}[Q_{\text{in}}(\Gamma) > J^{-1} \mid \mathcal{A}_J^\Theta(r, R)] > c. \quad (7.29)$$

And similar result holds for any configuration of inner faces with J faces. From this we can prove (by almost identical arguments) the counterpart of Proposition 7.17 in the plane case.

Proposition 7.20. *For any $K \geq j \geq 2$, there exists $c = c(j, K) > 0$ such that for any $10j \cdot 2^K \leq 2^K r \leq 2^K R_1 \leq R_2 \leq R$, let Γ be the set of interfaces crossing $A(R_1, R_2)$, then the following holds. For any configuration of outer faces Θ around C_{R_2} with no more than K faces, any color configuration ω_0 coincides with Θ and satisfies $\mathbb{P}[\mathcal{X}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$,*

$$\mathbb{P}[|\Gamma| = J \text{ and } Q_{\text{ex}}(\Gamma) > J^{-1} \mid \mathcal{X}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] > c. \quad (7.30)$$

Moreover, the same claim also holds for $\mathcal{Y}_j(r, R)$.

The first claim (7.30) can be deduced similarly as before. For the second claim, we note that although conditioning on $\mathcal{Y}_j(r, R)$ involves more requirements on connectedness, it could be resolved by similar gluing technique given in the proof of Lemma 3.3 in [8]. We omit the details.

7.3.7 Percolation exploration process and scaling limits

In this subsection, we define the percolation exploration process and discuss its power-law convergence towards SLE_6 obtained by Binder and Richards. In particular, we give in Proposition 7.22

a variant of their results tailored for our work. We will also show that the exploration process is stable under perturbations of the boundary in Proposition 7.23.

Given a Jordan domain Ω and two boundary points a, b , for any $0 < \eta < 1$, recall that Ω_η is a connected subgraph of \mathbb{T}_η (where we assume that η is small enough) and a_η and b_η are two vertices on the boundary of Ω_η . In addition, a_η and b_η divide the boundary of Ω_η into two connected parts, and we will assign red and blue colors to the external vertices adjacent to these two parts, respectively, according to the context. Define the exploration process γ_η as a directed path from a_η to b_η on \mathbb{T}_η^* which is the interface between the two red and blue clusters containing the boundary.

The Schramm-Loewner evolution (SLE), first introduced by Schramm in [21], is a family of conformally invariant random curves that serves as the canonical candidate for the scaling limit of curves, in particular interfaces of various critical models in 2D. In the case of critical planar percolation, Smirnov proves in [24] that the scaling limit of the exploration process γ_η is indeed the chordal SLE₆ path from a to b in Ω , which we denote by γ below.

As discussed in Section 7.1.2, Binder and Richards improve the above convergence by giving a power-law rate for curves up to some stopping time. We now describe their results in more detail. Given $U \subset \Omega$, let T_η be the first time that γ_η enters U_η and T be the first time that γ enters U . Define the distance d between two paths $l_1 : [0, t_1] \rightarrow \mathbb{R}^2$ and $l_2 : [0, t_2] \rightarrow \mathbb{R}^2$ as

$$d(l_1, l_2) = \inf_{\theta} \sup_{0 \leq s \leq t_1} |l_2(\theta(s)) - l_1(s)|,$$

where θ runs over all increasing homeomorphisms from $[0, t_1]$ to $[0, t_2]$. We call Ω a **nice domain** if Ω is a domain with a piece-wise smooth boundary. Note that the domains we consider in Sections 7.5.1 and 7.5.3 (both denoted by Ω) are nice domains.

In Theorem 4.1.11 of [20], it is proved that under some coupling of γ and γ_η , they are close upon specific stopping times with high probability. We collect their results here.

Theorem 7.21 ([20]). *When Ω is a nice domain and $U = B(b, \epsilon) \cap \Omega$ for some $\epsilon > 0$, there exists $u = u(\Omega, U) > 0$ such that under some coupling of γ and γ_η*

$$\mathbb{P} [d(\gamma_\eta|_{[0, T_\eta]}, \gamma|_{[0, T]}) > \eta^u] = O(\eta^u) \quad \text{as } \eta \rightarrow 0.$$

Combined with estimates on critical percolation, we can prove the following proposition which states that the time can be taken as the hitting time of a nice domain containing the endpoint in our setup.

Proposition 7.22. *For a nice domain Ω and some nice domain $U \subset \Omega$ that contains a neighbor of b , there exists $u = u(\Omega) > 0$ such that under some coupling of γ and γ_η (which we will later call **good coupling**),*

$$\mathbb{P} \left[d \left(\gamma_\eta|_{[0, T_\eta]}, \gamma|_{[0, T]} \right) > \eta^u \right] < O(\eta^u). \quad (7.31)$$

Proof. Fix $\epsilon > 0$ such that $U \cap \Omega \supset B(b, \epsilon) \cap \Omega$. We write T'_η for the first time that γ_η enters $B(b_\eta, \epsilon)$ and T' for the first time that γ enters $B(b, \epsilon)$. Note that the stopping times in Theorem 7.21 can be chosen as T'_η and T' . Hence, there exists $c > 0$ such that under some coupling \mathbb{P} of T_η and T ,

$$\mathbb{P} \left[d \left(\gamma_\eta|_{[0, T'_\eta]}, \gamma|_{[0, T']} \right) > \eta^c \right] < O(\eta^c). \quad (7.32)$$

We now show that \mathbb{P} is indeed a good coupling. Since $U_{[\eta]} \cap \Omega_{[\eta]} \supset B(b_\eta, \epsilon) \cap \Omega_{[\eta]}$ and $U \cap \Omega \supset B(b, \epsilon) \cap \Omega$, we have $T_\eta \leq T'_\eta$ and $T \leq T'$. Now, it is sufficient to prove that γ_η and γ respectively hit U_η and U almost at the same time and place with high probability. Write

$$\mathcal{F} = \left\{ \text{there does not exist } x \text{ in } \partial U_\eta \text{ such that } \gamma_\eta \text{ enters } B(x, \eta^c) \text{ but does not hit } U_\eta \right. \\ \left. \text{between the last entering and first leaving times of } B(x, \eta^{c'}), \text{ and the same for } \gamma \right\},$$

here c' is a constant to be chosen and $c > c' > 0$. If \mathcal{F}^c happens, we have a half-plane 3-arm event from a η^c -ball on $\partial U_{[\eta]}$ to distance $\eta^{c'}$ ¹¹, or, at a non-smooth point or a point η^c close to $\partial \Omega_{[\eta]}$, a 1-arm event in the half-plane squared. So, $\mathbb{P}[\mathcal{F}^c] \leq O(\eta^{-c}) \times O(\eta^{(1+\delta)(c-c')}) + O(1) \times O(\eta^{\delta(c-c')})$. Thus, we can choose c' small such that $\mathbb{P}[\mathcal{F}^c] \leq O(\eta^{c''})$. Next, we show that

$$\mathcal{F} \cap \left\{ d \left(\gamma_\eta|_{[0, T'_\eta]}, \gamma|_{[0, T']} \right) \leq \eta^c \right\} \subset \left\{ d \left(\gamma_\eta|_{[0, T_\eta]}, \gamma|_{[0, T]} \right) \leq \eta^{c'} \right\}.$$

Assume that $d \left(\gamma_\eta|_{[0, T'_\eta]}, \gamma|_{[0, T']} \right) \leq \eta^c$. When γ_η hits $U_{[\eta]}$, the process γ also gets η^c close to U . Then, γ will hit U before traveling a distance of $\eta^{c'}$, otherwise there will be a point x on ∂U such that γ enters $B(x, \eta^c)$ but does not hit U before leaving $B(x, \eta^{c'})$ which contradicts \mathcal{F} . Similarly, when γ hits U , the process γ_η will hit $U_{[\eta]}$ before leaving distance $\eta^{c'}$. So, γ_η and γ hit respectively

¹¹Careful readers will find that the boundary is not absolutely straight so these are not half-plane 3-arm events defined before and hence we cannot directly apply estimates from Remark 7.12. However, it is relatively easy to show that the probability that there exist polychromatic arms crossing a $(\pi + 0.01)$ -cone of radii R_1 and R_2 is also smaller than $(R_1/R_2)^{1+c}$. Another issue is that the boundary here can also be tilted, but similar arguments hold for a tilted $(\pi + 0.01)$ -cone as well thanks to the conformal invariance of the scaling limit. We will also meet similar issues later, but they can be handled in the same way.

$U_{[\eta]}$ and U almost at the same time and place, and their d -distance is smaller than $\eta^{c'}$. This completes the proof of the above relationship. Therefore, under this coupling

$$\mathbb{P} \left[d(\gamma_\eta|_{[0, T_\eta]}, \gamma|_{[0, T]}) > \eta^{c'} \right] \leq \mathbb{P} \left[d(\gamma_\eta|_{[0, T'_n]}, \gamma|_{[0, T']}) > \eta^c \right] + \mathbb{P}[\mathcal{F}^c] \leq O(\eta^c) + O(\eta^{c''}).$$

Letting $u = \min\{c, c', c''\}$ yields (7.31) as desired. \square

We will also use a variant of this proposition, i.e., Proposition 7.35, in the proof of Proposition 7.4.

The next proposition states that the exploration process is insensitive to the change of boundaries. More precisely, when the boundary changes by η^δ , then the exploration process changes by at most $\eta^{c(\delta)}$ with probability at least $(1 - \eta^{c(\delta)})$.

Proposition 7.23. *Suppose Ω is a nice domain and $\Omega_{[\eta]}$ be the discretization of Ω defined in Section 7.3.1. Let Ω'_η be a simply connected sub-graph of $\mathbb{T}_{[\eta]}^*$ such that if u, v are two vertices of Ω'_η then the edge uv also belongs to the edge set of $\Omega_{[\eta]}$. Let a'_η, b'_η be two boundary points of Ω'_η . If for some $\delta > 0$,*

$$d(\partial\Omega_{[\eta]}, \partial\Omega'_\eta) < \eta^\delta, \quad d(a_\eta, a'_\eta) < \eta^\delta, \quad d(b_\eta, b'_\eta) < \eta^\delta.$$

then there exists a coupling \mathbb{P} of γ_η and γ'_η and a constant $c(\delta) > 0$ such that

$$\mathbb{P} \left[d(\gamma_\eta, \gamma'_\eta) > \eta^c \right] < O(\eta^c).$$

Proof. Without loss of generality, we assume that $\Omega'_\eta \supset \Omega_{[\eta]}$ (otherwise we can compare $\Omega_{[\eta]}$ and Ω'_η both with Ω''_η , where $\Omega''_\eta = \{x \in \Omega_{[\eta]} : d(x, \partial\Omega_{[\eta]}) > \eta^\delta\}$). We can couple the critical percolation on $\Omega_{[\eta]}$ and Ω'_η such that the two configurations have the same color on $\Omega_{[\eta]}$. Under this coupling, the exploration processes in these two domains can only be different after one of them hits $\partial\Omega_{[\eta]}$ (since they have the same path in the interior of $\Omega_{[\eta]}$). Furthermore, we can find some $c \in (0, \delta)$ such that with more than $(1 - \eta^c)$ probability, each time they hit $\Omega_{[\eta]}$, they will remain the same after η^c distance (otherwise we have a half plane 3-arm event from an η^δ ball on the boundary to distance η^c , or, not a non-smooth point, a 1-arm event in the whole plane. This happens with probability less than η^c when c is close enough to 0). This completes the proof. \square

Remark 7.24. With similar arguments, we can show that the sharp asymptotics we obtain in this work for arm events under the specific discretization, also hold for other discretizations, as they

differ by at most a power of the mesh size. In other words, arm events are insensitive to the way we discretize domains. In particular, the good coupling we construct in Proposition 7.22 (as well as Proposition 7.35) also works for the discretized domains under the convention given at the end of Section 7.3.1.

7.4 Couplings and conditional arm probabilities

In this section, we will give various couplings concerning arm events in the half-plane and the plane; see Propositions 7.25, 7.27 and 7.30 and deduce useful estimates on the conditional arm probabilities from these couplings; see Propositions 7.28, 7.29, 7.32, 7.34. In [8], the authors have established the coupling results concerning one-arm ([8, Proposition 5.2]) and four-arm ([8, Proposition 3.1]) events in the plane, respectively. Although many techniques developed in [8] can be adapted for variants of arm events we are considering here (e.g., $\mathcal{H}_j(r, R)$, $\mathcal{X}_j(r, R)$ and $\mathcal{Y}_j(r, R)$), certain intricacy arises from our choice of such events, causing extra difficulties. We postpone the proofs to Section 7.6 as they are rather technical and independent of the main story.

7.4.1 The half-plane case

In this subsection, we deal the coupling results in the half-plane. For simplicity of presentation, the $j = 1$ case and the $j \geq 2$ cases are stated separately.

One-arm In this paragraph, we will state the coupling result for one-arm event in the half-plane. We say that a b-path in the half-plane is a **circuit** if it connects the discretized negative real line to the discretized positive real line. If $\eta < r < u$, let $\Gamma_{\text{out}}(r, u)$ (resp. $\Gamma_{\text{in}}(r, u)$) be the outermost (resp. innermost) red circuit in the semi-annulus $A^+(r, u)$, i.e., the red circuit which is closest to C_u^+ (resp. C_r^+). If such circuits do not exist, then set $\Gamma(r, u) = \emptyset$.

Proposition 7.25. *There exists $\delta > 0$ such that for all $100 < 10r < R$ and $m \in (1.1, 10)$, denoting $u := \sqrt{rR}$, then the following hold:*

- (Inward coupling) *There is a coupling \mathbb{Q} of the conditional laws $\mathbb{P}[\cdot \mid \mathcal{H}_1(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{H}_1(r, mR)]$ such that if we sample $(\omega_1, \omega_2) \sim \mathbb{Q}$, then with probability at least $(1 - (r/R)^\delta)$, $\Gamma_{\text{out}}(r, u)$ for both ω_1 and ω_2 are non-empty and identical. Furthermore, the percolation con-*

figurations in the connected component of $\mathbb{H} \setminus \Gamma_{\text{out}}(r, u)$ whose boundary contains 0 are also identical.

- (Outward coupling) A similar coupling exists for the conditional laws $\mathbb{P}[\cdot \mid \mathcal{H}_1(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{B}_1(r, R)]$ with Γ_{out} and 0 replaced by Γ_{in} and ∞ respectively.

Remark 7.26. In the statement above, we take $u = \sqrt{rR}$ merely for convenience. In fact, it can be taken as $r^d R^{1-d}$ for any $0 < d < 1$ (with the corresponding $\delta(d)$). The same applies to all couplings in this section.

The proof of the above proposition is essentially the same as that of [8, Proposition 5.2], which deals with one-arm events in the plane in which the key tools are the RSW theory and FKG inequality. Therefore, we omit the proof and refer the readers to the said paper for details.

j -arms with $j \geq 2$ In this paragraph, we concentrate on the j -arm events with $j \geq 2$ in the half-plane. We shall present couplings in both the inward and outward directions simultaneously for they share the same virtue. Note that for $j \geq 2$, $\mathcal{H}_j(r, R)$ is equivalent to the event that there are at least $(j-1)$ interfaces starting from C_R^+ and end at their first hitting of $[-r, r]$; similarly, $\mathcal{B}_j(r, R)$ is equivalent to the event that there are at least $(j-1)$ interfaces crossing the annulus $A^+(r, R)$. In both cases, we write $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ (where $n \geq j-1$) for the set of interfaces crossing the whole region, and order them counterclockwise. We say two interfaces in Γ are **adjacent** if there is a hexagon on \mathbb{T}^* touched by both of them.

Proposition 7.27. For any $j \geq 2$, there exists $\delta(j) > 0$ such that for any $100j < 10r < R$ and $m \in (1.1, 10)$, denoting $u = \sqrt{rR}$, the following hold:

- (Inward coupling) There is a coupling \mathbb{Q} of $\mathbb{P}[\cdot \mid \mathcal{H}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{H}_j(r, mR)]$ such that, if we sample $(\omega_1, \omega_2) \sim \mathbb{Q}$, then with probability at least $(1 - (r/R)^\delta)$, there exists a common configuration of outer faces Θ^* with j faces around C_u^+ in both ω_1 and ω_2 , and ω_1 coincides with ω_2 in \mathcal{V}_{Θ^*} . Furthermore, when this is the case, Θ^* is a stopping set (recall the definition in Section 7.3.4) and for any $\eta < r' < r$, we have

$$\mathbb{P}[\mathcal{H}_j(r', R) \mid \mathcal{H}_j(r, R), \Theta^*] = \mathbb{P}[\mathcal{H}_j(r', mR) \mid \mathcal{H}_j(r, mR), \Theta^*]. \quad (7.33)$$

- (Outward coupling) A similar coupling exists for $\mathbb{P}[\cdot \mid \mathcal{H}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{B}_j(r, R)]$ with outer faces replaced by inner faces. For any $R' > R$, when the coupling succeeds (so that the

common inner faces Θ^* exists), we have

$$\mathbb{P}[\mathcal{H}_j(r, R') \mid \mathcal{H}_j(r, R), \Theta^*] = \mathbb{P}[\mathcal{B}_j(r, R') \mid \mathcal{B}_j(r, R), \Theta^*]. \quad (7.34)$$

As a consequence of these couplings in Proposition 7.25 (for the $j = 1$ case) and Proposition 7.27 (for the $j \geq 2$ cases), we know that the law $\mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, n)]$ will coincide with $\mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, mn)]$ at scale n^α with high probability for $\alpha \in (0, 1)$ ¹². From the (slightly stronger) separation lemma Proposition 7.17, we can show that the failure event will not contribute much to $\mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n)]$ much. Thus, we obtain the following comparison of conditional arm probabilities.

Proposition 7.28. *For any $r \geq r_h(j)$, $m \in (1.1, 10)$ and $\alpha \in (0, 1)$,*

$$\mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n)] = \mathbb{P}[\mathcal{H}_j(r, mn) \mid \mathcal{H}_j(n^\alpha, mn)] (1 + O(n^{-c})). \quad (7.35)$$

Here, $O(n^{-c})$ may depend on r and α , but not m .

Proof. We only give the proof for $j \geq 2$, for the case $j = 1$ is similar (and easier indeed). Denote $\beta = (1 + \alpha)/2$, and choose a large integer $K_0 = K_0(j, \alpha)$ such that

$$\mathbb{P}[\mathcal{B}_{K_0}(n^\alpha, n^\beta)] \leq n^{-\beta_j - 1}. \quad (7.36)$$

Consider probability measures $\mathbb{P}_1 = \mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, n)]$ and $\mathbb{P}_2 = \mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, mn)]$ with sample spaces Ω_1 and Ω_2 . For a pair $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$, we say $\omega_1 =_{n^\beta} \omega_2$ if there exists a common configuration Θ^* of outer faces around $C_{n^\beta}^+$, such that ω_1 and ω_2 are identical in \mathcal{V}_{Θ^*} . Divide $\Omega_1 \times \Omega_2$ into the union of $\mathcal{K}_1 = \{(\omega_1, \omega_2) : \omega_1 =_{n^\beta} \omega_2\}$, $\mathcal{K}_2 = \{(\omega_1, \omega_2) : \omega_1 \in \mathcal{B}_{K_0}(n^\alpha, n^\beta) \text{ or } \omega_2 \in \mathcal{B}_{K_0}(n^\alpha, n^\beta)\} \setminus \mathcal{K}_1$ and $\mathcal{K}_3 = (\mathcal{K}_1 \cup \mathcal{K}_2)^c$. By Proposition 7.27, we can couple $\mathbb{P}_1, \mathbb{P}_2$ by \mathbb{Q} in a way that

$$\mathbb{Q}[\mathcal{K}_1] \geq 1 - n^{-C} \text{ for some } C(\alpha, j) > 0. \quad (7.37)$$

We define a pair of random stopping sets (Θ_1, Θ_2) under \mathbb{Q} as follow. For $(\omega_1, \omega_2) \sim \mathbb{Q}$, since Θ^* is a stopping set by Proposition 7.27, we can perform an exploration process outside $C_{n^\beta}^+$ for both configurations to check that whether there exists a common configuration of outer faces Θ^* with j faces for both $\omega_i, i = 1, 2$ and without exploring any hexagon in \mathcal{V}_{Θ^*} when this is the case. If so (see the left part of Figure 7.6), we then further explore the $(j - 1)$ interfaces starting from the end-points of Θ^* and end upon reaching $C_{n^\alpha}^+$ and hence obtain a pair of configurations of outer

¹²Not to be confused with the arm exponents α_j . Same for β in the proof of Proposition 7.28 below.

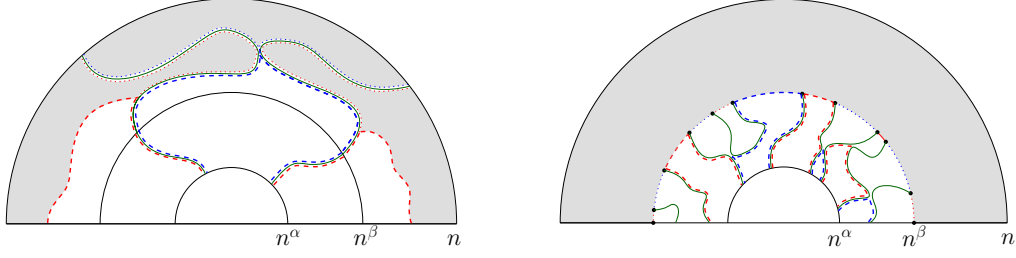


Figure 7.6: Proof of Proposition 7.28. Sketch for $\mathbb{P}[\cdot \mid \mathcal{H}_3(n^\alpha, n)]$. **Left:** The case when the coupling succeeds: there is a common configuration of outer faces Θ^* with 3 faces around $C_{n^\beta}^+$ for both $\omega_1 \sim \mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, n)]$ and $\omega_2 \sim \mathbb{P}[\cdot \mid \mathcal{H}_j(n^\alpha, mn)]$. The two interfaces that cross the annulus $A^+(n^\alpha, n)$ are in green. Faces Θ^* together with hexagons that are adjacent to these two interfaces further create a common configuration of outer faces around $C_{n^\alpha}^+$, which is sketched in dashed red/blue curves. The region that has been explored (\mathcal{V}_{Θ^*}) is in gray. **Right:** The case when the coupling fails. In this case, we reveal the color of all hexagons in $A^+(n^\beta, n)$, which is illustrated in gray. Then, the color of all hexagons neighboring $C_{n^\beta}^+$ are revealed, which is sketched in dotted/dashed red/blue arcs on $C_{n^\beta}^+$. Running all of the exploration processes (in green) from $C_{n^\beta}^+$ towards $C_{n^\alpha}^+$, we obtain a configuration of outer faces around $C_{n^\alpha}^+$, which is sketched in dashed red/blue curves. In this case, the number of faces could be larger than j .

faces $(\Theta_1(\omega_1), \Theta_2(\omega_2))$ around $C_{n^\alpha}^+$. Otherwise (see the right part of Figure 7.6), we further explore all the hexagons outside $C_{n^\beta}^+$ together with all the interfaces in $A^+(n^\alpha, n^\beta)$ from $C_{n^\beta}^+$ to $C_{n^\alpha}^+$ for both $\omega_i, i = 1, 2$. This will also yield a pair of configurations of outer faces $(\Theta_1(\omega_1), \Theta_2(\omega_2))$ around $C_{n^\alpha}^+$. Note that in either case, all hexagons in $\mathcal{V}_{\Theta_i(\omega_i)}, i = 1, 2$ are left unexplored, hence we get a pair of stopping sets $(\Theta_1, \Theta_2)(\omega_1, \omega_2) = (\Theta_1(\omega_1), \Theta_2(\omega_2))$.

For $(\omega_1, \omega_2) \in \mathcal{K}_1$, $\Theta_1(\omega_1)$ and $\Theta_2(\omega_2)$ are identical with j faces, and we deduce from (7.33) that

$$\mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n), \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1] = \mathbb{P}[\mathcal{H}_j(r, mn) \mid \mathcal{H}_j(n^\alpha, mn), \omega_{\mathcal{D}_{\Theta_2(\omega_2)}} = \omega_2]. \quad (7.38)$$

For $(\omega_1, \omega_2) \in K_3$, we claim it holds uniformly that

$$\mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n), \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1] \leq n^{-\beta_j \alpha + o(1)}, \quad (7.39)$$

$$\mathbb{P}[\mathcal{H}_j(r, mn) \mid \mathcal{H}_j(n^\alpha, mn), \omega_{\mathcal{D}_{\Theta_2(\omega_2)}} = \omega_2] \leq n^{-\beta_j \alpha + o(1)}. \quad (7.40)$$

Since $\mathcal{H}_j(r, n) \subset \mathcal{B}_j(r, 2^{-K_0} n^\alpha) \cap \mathcal{B}_j(2^{-K_0} n^\alpha, n)$ and $\mathcal{B}_j(r, 2^{-K_0} n^\alpha)$ is independent with $\mathcal{B}_j(2^{-K_0} n^\alpha, n)$ and $\{\omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1\}$, then uniformly for $(\omega_1, \omega_2) \in K_3$, the left hand side of (7.39) is bounded by

$$\frac{\mathbb{P}[\mathcal{B}_j(r, 2^{-K_0} n^\alpha)] \mathbb{P}[\mathcal{B}_j(2^{-K_0} n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1]} \leq n^{-\beta_j \alpha + o(1)} \cdot \frac{\mathbb{P}[\mathcal{B}_j(2^{-K_0} n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1]}.$$

Noting that $\Theta_1(\omega_1)$ has no more than K_0 faces for any $(\omega, \omega') \in K_3$, and letting \mathcal{U} be the event that the interfaces crossing $A^+(2^{-K_0} n^\alpha, n^\alpha)$ are well-separated on $C_{2^{-K_0} n^\alpha}^+$, it follows that

$$\begin{aligned} \mathbb{P}[\mathcal{H}_j(n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1] &\stackrel{\text{RSW}}{\geq} c(j, K_0) \mathbb{P}[\mathcal{U} \cap \mathcal{B}_j(2^{-K_0} n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1] \\ &\stackrel{(7.23)}{\geq} c'(j, K_0) \mathbb{P}[\mathcal{B}_j(2^{-K_0} n^\alpha, n) \mid \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1]. \end{aligned}$$

This proves (7.39), and (7.40) follows in the same fashion.

Now, by Bayesian formula, writing

$$F := \mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n), \omega_{\mathcal{D}_{\Theta_1(\omega_1)}} = \omega_1], \quad G := \mathbb{P}[\mathcal{H}_j(r, mn) \mid \mathcal{H}_j(n^\alpha, mn), \omega_{\mathcal{D}_{\Theta_2(\omega_2)}} = \omega_2]$$

for short, we have

$$\left| \mathbb{P}_1[\mathcal{H}_j(r, n)] - \mathbb{P}_2[\mathcal{H}_j(r, mn)] \right| = \left| \sum_{(\Theta_1, \Theta_2)} (F - G) \mathbb{Q}[(\Theta_1, \Theta_2) = (\Theta_1, \Theta_2)] \right|. \quad (7.41)$$

The RHS of (7.41) can be bounded by

$$\left| \sum_{(\Theta_1, \Theta_2)} (F - G) \mathbb{Q}[(\Theta_1, \Theta_2) = (\Theta_1, \Theta_2), \mathcal{K}_1] \right| + \sum_{i=2}^3 \sum_{(\Theta_1, \Theta_2)} (F + G) \mathbb{Q}[(\Theta_1, \Theta_2) = (\Theta_1, \Theta_2), \mathcal{K}_i].$$

The first term equals 0 by (7.38), the second term (that corresponds to \mathcal{K}_2 in the sum) is bounded by $2\mathbb{Q}[\mathcal{K}_2] \leq 4n^{-\beta_j - 1}$ from (7.36), and the last term is bounded by $n^{-\beta_j \alpha - C + o(1)}$ from (7.37), (7.39) and (7.40). As a result we see $|\mathbb{P}_1[\mathcal{H}_j(r, n)] - \mathbb{P}_2[\mathcal{H}_j(r, n)]| \leq n^{-\beta_j \alpha - C + o(1)}$ for some $C > 0$. In addition, by Lemma 7.11 and (7.17) in Remark 7.12, $\mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(n^\alpha, n)] = n^{-\beta_j \alpha + o(1)}$. The claim (7.35) hence follows. \square

Proposition 7.29. *For any $r \geq r_h(j) (\geq r_b(j))$, $m \in (1.1, 10)$ and $\alpha \in (0, 1)$,*

$$\mathbb{P}[\mathcal{B}_j(r, n) \mid \mathcal{B}_j(r, n^\alpha)] = \mathbb{P}[\mathcal{H}_j(r, n) \mid \mathcal{H}_j(r, n^\alpha)] \left(1 + O(n^{-c})\right). \quad (7.42)$$

Proof. In contrast to Proposition 7.28, this is an outward coupling, in which case we do not need to spare an additional scale n^β to use Proposition 7.17, and thus much easier to deal with. Concretely, we first examine whether $\mathbb{P}[\cdot \mid \mathcal{B}_j(r, n^\alpha)]$ and $\mathbb{P}[\cdot \mid \mathcal{H}_j(r, n^\alpha)]$ induce the same configuration of inner faces around $C_{n^\alpha}^+$, which will happen with probability larger than $(1 - n^{-C})$ by (7.34). If it fails, for both events $\mathcal{B}_j(r, n)$ and $\mathcal{H}_j(r, n)$, we further need to fulfill an arm event $\mathcal{B}_j(n^\alpha, n)$ which has probability bounded by $n^{-\beta_j \alpha + o(1)}$ by (7.17). The result follows immediately. \square

7.4.2 The plane cases

In this subsection, we provide the coupling results in the plane, which are analogous to those in the half plane in the previous subsection. The one-arm and four-arm cases have already been developed in [8], but as new difficulties arise for the variants we consider in this work and for general j , we decide to include a proof, which we postpone to the last section. Recall the definitions of $\mathcal{X}_j, \mathcal{Y}_j, \mathcal{A}_j$ for $j \geq 2$ in (7.15) and (7.14) and below (7.15) resp. We let

$$J = 2\lfloor j/2 \rfloor.$$

In the next proposition, we couple together conditional laws in the plane case. For discussions of these couplings, see Proposition 7.31 below.

Proposition 7.30. *For any $j \geq 2$, there exists $\delta(j) > 0$ such that for all $100j\eta \leq 10r \leq R$ and $m \in (1.1, 10)$, denoting $u = \sqrt{rR}$, then the following hold.*

- (Inward coupling) *There is a coupling \mathbb{Q} of the conditional laws $\mathbb{P}[\cdot \mid \mathcal{X}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{A}_j(r, R)]$ such that if we sample $(\omega_1, \omega_2) \sim \mathbb{Q}$, then with probability at least $(1 - (r/R)^\delta)$, there exists a common configuration of outer faces Θ^* with J faces around C_u in both ω_1 and ω_2 , and ω_1 coincides with ω_2 in \mathcal{V}_{Θ^*} . Furthermore, when this is the case, Θ^* is a stopping set and for any $1 \leq r' \leq r$,*

$$\mathbb{P}[\mathcal{X}_j(r', R) \mid \mathcal{X}_j(r, R), \omega_{\mathcal{D}_{\Theta^*}} = \omega_1] = \mathbb{P}[\mathcal{A}_j(r', R) \mid \mathcal{A}_j(r, R), \omega_{\mathcal{D}_{\Theta^*}} = \omega_2]. \quad (7.43)$$

- (Another inward coupling) *A similar coupling exists for the conditional laws $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR)]$. In this case, if the coupling succeeds (i.e., Θ^* exists),*

$$\mathbb{P}[\mathcal{Y}_j(r', R) \mid \mathcal{Y}_j(r, R), \omega_{\mathcal{D}_{\Theta^*}} = \omega_1] = \mathbb{P}[\mathcal{Y}_j(r', mR) \mid \mathcal{Y}_j(r, mR), \omega_{\mathcal{D}_{\Theta^*}} = \omega_2]. \quad (7.44)$$

Remark 7.31. The major difference of these couplings with Proposition 7.27 lies in (7.44). In the half-plane case, the “domain Markov property” can be easily applied, and thus (7.33) and (7.34) follows naturally from the existence of common faces for both percolation configurations. In the plane case, due to the complicity of the events $\mathcal{X}_j(r, R)$ and $\mathcal{Y}_j(r, R)$, (7.44) does not hold trivially. We tackle this kind of difficulty in Section 7.6 by coupling together extra structures rather than just configurations of faces.

As a result of Proposition 7.30, we get the following estimates on conditional arm probabilities.

Proposition 7.32. *For any $r \geq r_y(j)$, $m \in (1.1, 10)$ and $\alpha \in (0, 1)$, we have*

$$\mathbb{P}[\mathcal{Y}_j(r, n) | \mathcal{Y}_j(n^\alpha, n)] = \mathbb{P}[\mathcal{Y}_j(r, mn) | \mathcal{Y}_j(n^\alpha, mn)] \left(1 + O(n^{-c})\right), \quad (7.45)$$

where $O(n^{-c})$ may depend on r and α , but not m .

Proposition 7.33. *For any $r \geq r_x(j)(= r_a(j))$ and $\varepsilon > 0$, we have*

$$\mathbb{P}[\mathcal{X}_j(r, n) | \mathcal{X}_j(\varepsilon n, n)] = \mathbb{P}[\mathcal{A}_j(r, n) | \mathcal{A}_j(\varepsilon n, n)] \left(1 + O(\varepsilon^c)\right). \quad (7.46)$$

The proof of Proposition 7.32 is almost identical to that of Proposition 7.28, and by similar arguments and replacing n^α with εn , we obtain a similar proof of Proposition 7.33. We omit the proofs.

We now consider outward coupling in the plane. Recall the definition of $\mathcal{Z}_j(r, R)$. There are similar couplings of $\mathbb{P}[\cdot | \mathcal{A}_j(r, n^\alpha)]$ and $\mathbb{P}[\cdot | \mathcal{Z}_j(r, n^\alpha)]$ as in the half-plane case, from which we can deduce a comparison of the probability of \mathcal{X}_j conditioned on \mathcal{A}_j and \mathcal{Y}_j on \mathcal{Z}_j , which is summarized in the following proposition. As the key idea is essentially same as before, for brevity we choose not to state the precise coupling results, and only give a sketch of the proof for the comparison result.

Proposition 7.34. *For any $r \geq r_z(j)(\geq r_a(j))$, $m \in (1.1, 10)$ and $\alpha \in (0, 1)$, we have*

$$\mathbb{P}[\mathcal{X}_j(r, n) | \mathcal{A}_j(r, n^\alpha)] = \mathbb{P}[\mathcal{Y}_j(r, n) | \mathcal{Z}_j(r, n^\alpha)] \left(D(j) + O(n^{-c})\right) \quad (7.47)$$

where $D(j)$ is a constant given by

$$D(j) = \begin{cases} 1, & j \equiv 1 \pmod{2} \\ \frac{j}{2}, & j \equiv 2 \pmod{4} \\ \frac{j}{4}, & j \equiv 0 \pmod{4} \end{cases} \quad (7.48)$$

and $O(n^{-c})$ may depend on r , α and j , but not m .

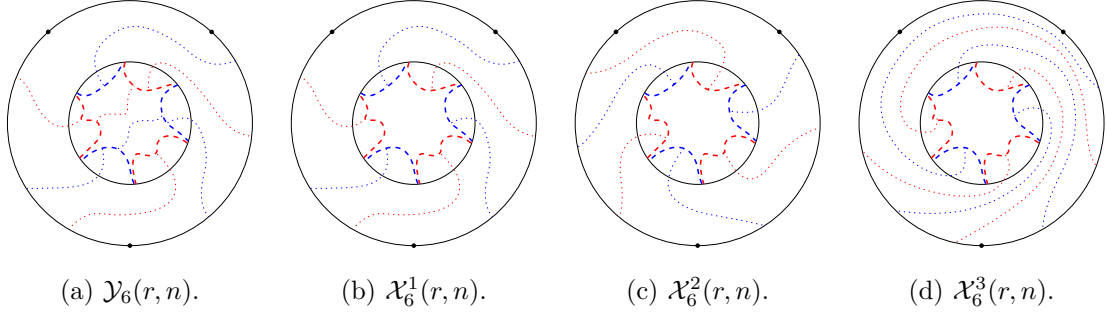


Figure 7.7: An illustration of the events $\mathcal{Y}_6(r, n)$ and $\mathcal{X}_6(r, n)$ of three different types of connecting patterns. All four annuli are $A(n^\alpha, n)$ with the same configuration of inner faces Θ^* in dashed red/blue curves. The three marked points on C_n are $n \cdot a$, $n \cdot b$ and $n \cdot c$, respectively (from bottom, counterclockwise). The outer connecting patterns for Θ^* are sketched in dotted red/blue curves, and there is an additional inner connecting pattern in the first picture which results in the only possibility for the outer connection for $\mathcal{Y}_6(r, n)$.

We first explain how the constant $D(j)$ in (7.48) appears. As in Proposition 7.30, we can couple $\mathbb{P}[\cdot \mid \mathcal{A}_j(r, n^\alpha)]$ and $\mathbb{P}[\cdot \mid \mathcal{Z}_j(r, n^\alpha)]$ in a way such that if we sample (ω_1, ω_2) according to this coupling, ω_1 and ω_2 will coincide on a configuration of inner faces Θ^* around C_{n^α} with high probability. We now consider the events $\mathcal{X}_j(r, n)$ and $\mathcal{Y}_j(r, n)$ conditioned on the color configuration in \mathcal{D}_{Θ^*} . When j is odd, conditioned on \mathcal{D}_{Θ^*} , $\mathcal{X}_j(r, n)$ is equivalent to $\mathcal{Y}_j(r, n)$ and the proof is almost identical with the previous cases. However, when j is even, these two events are distinct: there are $j/2$ kinds of connecting patterns for $\mathcal{X}_j(r, n)$, since the information inside Θ^* cannot tell us which faces should be connected to the corresponding segments on C_n ; but for $\mathcal{Y}_j(r, n)$, there are only 1 (if $j \equiv 2 \pmod{4}$) or 2 (if $j \equiv 0 \pmod{4}$) possible connecting patterns, because the information inside Θ^* already determines the order of faces in Θ^* . See Figure 7.7 for an illustration of the case when $j = 6$. If we realize that all possible connecting patterns have nearly the same probability thanks to the coupling results, then the definition of $D(j)$ in (7.48) makes sense.

Sketch of Proof for Proposition 7.34 when j is even. Denote \mathcal{G} for the event that there are exactly j interfaces crossing $A(r, n)$, then comparing with $\mathcal{X}_j(r, n)$ and $\mathcal{Y}_j(r, n)$, \mathcal{G}^c has negligible probability by BK-Reimer's inequality. We partition $\mathcal{X}_j(r, n) \cap \mathcal{G}$ into the disjoint union of $D(j)$ events according

to the locations of interfaces. More precisely, let $\hat{\Gamma}$ be the set of the $\frac{j}{2}$ interfaces with red on their left (seen from outside to inside). We label each $\gamma \in \hat{\Gamma}$ in the following two ways: label γ by γ_i (resp. γ^i) with $1 \leq i \leq \frac{j}{2}$, if γ is the i -th element in $\hat{\Gamma}$ when counting around the circle C_r (resp. C_n) counterclockwise starting from $(0, -r)$ (resp. $(0, -n)$). For $1 \leq i \leq D(j)$, define $\mathcal{X}_j^i(r, n)$ as

$$\mathcal{X}_j^i(r, n) = \begin{cases} \mathcal{X}_j(r, n) \cap \mathcal{G} \cap \{\gamma_1 = \gamma^i\} & j \equiv 2 \pmod{4}, \\ \mathcal{X}_j(r, n) \cap \mathcal{G} \cap \{\gamma_1 = \gamma^i \text{ or } \gamma^{i+\frac{j}{4}}\} & j \equiv 0 \pmod{4}. \end{cases} \quad (7.49)$$

See Figure 7.7 for illustration of the events $\mathcal{X}_j^i(r, n)$'s. Then $\mathcal{X}_j(r, n) \cap \mathcal{G}$ is the disjoint union of $\mathcal{X}_j^i(r, n)$, $1 \leq i \leq D(j)$. We construct a coupling of $\mathbb{P}[\cdot \mid \mathcal{A}(r, n^\alpha)]$ and $\mathbb{P}[\cdot \mid \mathcal{Z}_j(r, n^\alpha)]$ for each $1 \leq i \leq D(j)$. In such couplings, we keep a record of the relative location of interfaces, so once we successfully couple a face, the conditional version of $\mathcal{X}_j^i(r, n)$ and $\mathcal{Y}_j(r, n) \cap \mathcal{G}$ are identical. This allows us to deduce that

$$\mathbb{P}[\mathcal{X}_j^i(r, n) \mid \mathcal{A}_j(r, n^\alpha)] = \mathbb{P}[\mathcal{Y}_j(r, n) \cap \mathcal{G} \mid \mathcal{Z}_j(r, n^\alpha)](1 + O(n^{-c}))$$

holds for each $1 \leq i \leq D(j)$, and the desired result follows readily. \square

7.5 Proof of main theorems

In this section, we will combine the coupling results from the previous section and the power-law convergence of the exploration process discussed in Section 7.3.7 to prove the main results of this work, namely Theorems 7.1, 7.2 and 7.3. In Subsections 7.5.1 and 7.5.2, we will prove Theorem 7.1, which is derived directly from Proposition 7.4. In order to prove Proposition 7.4, we use Propositions 7.28, 7.29 and 7.36 as inputs. In Subsections 7.5.3 and 7.5.4, we will prove Theorems 7.2 and 7.3. The former is derived directly from Proposition 7.5, while the latter theorem is a quick corollary of the former. In order to prove Proposition 7.5, we use Propositions 7.32 and 7.40 (the latter is an analogue of Proposition 7.36 in the plane), as inputs.

Note that in Sections 7.5.1 and 7.5.3, we will consider the rescaled lattice and the discretization scheme in Section 7.3.1 comes into play.

7.5.1 Comparison estimates in the half-plane

In this subsection, our main goal is Proposition 7.36 in which we compare the probabilities of \mathcal{H}_j , the modified half-plane arm events (recall (7.12) for definitions), at different scales.

As discussed in Section 7.1.2, the key ingredient is the convergence rate of percolation exploration process towards SLE_6 upon some stopping time as shown in Theorem 4.1.10 in [20] and discussed in Section 7.3.7. Instead of citing Proposition 7.22 directly, we will use a modified version stated as follows.

Recall the notation from Section 7.3.7. We consider the nice domain (in the sense of Section 7.3.1)

$$\Omega = B_1^+ \cup B^- \left((-3/4, 0), 1/4 \right), a = (1, 0) \text{ and } b = (-3/4, -1/4);$$

see Figure 7.8 for an illustration. We assign red color to the counterclockwise arc \widehat{ab} and blue color to the counterclockwise arc \widehat{ba} . Given mesh size $\eta > 0$, the percolation exploration process starts from a_η and end at b_η . Given $\alpha \in (0, 1)$ and $m > 0$, let $\tilde{T}_{1/n}$ (resp. $\tilde{T}_{1/(mn)}, \tilde{T}$) denote the first time that $\gamma_{1/n}$ (resp. $\gamma_{1/(mn)}, \gamma$) hits $[-1, -n^{\alpha-1}]$.

Recall (7.12) that $\mathcal{H}_j(r, R)$ is the event that in B_R^+ with the lattice scale $\eta = 1$ there exist j disjoint arms connecting $[-r, r]$ to C_R^+ with alternating colors. Rescaling the lattice by $\eta = 1/R$,¹³ we get that $\mathcal{H}_j(r, R)$ is equivalent to the event that in B_1^+ with the lattice scale η , there exist j such arms connecting $[-r/R, r/R]$ to C_1^+ . From here till the end of this subsection, with slight abuse of notation, we will tacitly assume the equivalence between events on the rescaled and those on the unrescaled lattices.

Proposition 7.35. *There exists $u(\Omega) > 0$ such that for any $\alpha \in (0, 1)$ there exists a coupling of $\gamma_{1/n}$ and γ (which we will refer to as the good coupling)*

$$\mathbb{P} \left[d \left(\gamma_{1/n} \big|_{[0, \tilde{T}_{1/n}]}, \gamma \big|_{[0, \tilde{T}]} \right) > n^{-u} \right] = O(n^{-u}).$$

One can further check that $O(n^{-u})$ is independent of $\alpha \in (0, 1)$.

Proof. The proof is the same as that of Proposition 7.22 except that here we let \mathcal{F} denote the event that there does not exist x in $[-1, -n^{\alpha-1}]$ such that $\gamma_{1/n}$ enters $B(x, n^{-c})$ ($c(\Omega)$ is the constant associated with the coupling in Theorem 4.1.10 of [20]) but does not hit $[-1, -n^{\alpha-1}]$ between the last entrance and first exit times of $B(x, n^{-c'})$, and the same for γ . We can still choose c' small such that $\mathbb{P}[\mathcal{F}^c] \leq O(n^{-c''})$ and show that under good coupling of $\gamma_{1/n}$ and γ

$$\mathcal{F} \cap \left\{ \left(\gamma_{1/n} \big|_{[0, T'_{1/n}]}, \gamma \big|_{[0, T']} \right) \leq n^{-c} \right\} \subset \left\{ d \left(\gamma_{1/n} \big|_{[0, \tilde{T}_{1/n}]}, \gamma \big|_{[0, \tilde{T}]} \right) \leq n^{-c'} \right\}.$$

¹³In fact we will consider the cases $R = n$ and $R = mn$ in this and the next subsection.

Hence,

$$\mathbb{P} \left[d \left(\gamma_{1/n} \big|_{[0, \tilde{T}_{1/n}]}, \gamma \big|_{[0, \tilde{T}]} \right) > n^{-c'} \right] \leq \mathbb{P} \left[d \left(\gamma_{1/n} \big|_{[0, T'_{1/n}]}, \gamma \big|_{[0, T']} \right) > n^{-c} \right] + \mathbb{P}[\mathcal{F}^c] \leq O(n^{-c}) + O(n^{-c''}).$$

Taking $u = \min\{c, c', c''\}$, we obtain the proposition. \square

Similarly, for the same constant u , we can show that one can further couple $\gamma_{1/mn}$ and γ (we still denote the law by \mathbb{P} , view it as the law of $\gamma_{1/n}$, $\gamma_{1/mn}$ and γ coupled together and call it the good coupling) such that

$$\mathbb{P} \left[d \left(\gamma_{1/(mn)} \big|_{[0, \tilde{T}_{1/(mn)}]}, \gamma \big|_{[0, \tilde{T}]} \right) > n^{-u} \right] = O(n^{-u}). \quad (7.50)$$

One can further check that $O(n^{-u})$ is independent of $\alpha \in (0, 1)$ and m in a finite interval away from zero, e.g., $(1.1, 10)$.

Recall the definition of $h_j(r, R)$ in (7.16). The following proposition states that h_j are almost the same at different scales within a power-law error at different scales when r is not too small compared to R .

Proposition 7.36. *There exists $c_8 > 0$ such that for all $\alpha \in (1 - c_8, 1)$ and $m \in (1.1, 10)$,*

$$h_j(n^\alpha, n) = h_j(mn^\alpha, mn) \left(1 + O(n^{-c}) \right).$$

Importantly, here $O(n^{-c})$ is independent of m and α !

To prove it, we need Lemmas 7.37 and 7.38. In Lemma 7.37, we will show that $\mathcal{H}_j(n^\alpha, n)$ can be determined by the exploration process $\gamma_{1/n}$ upon time $\tilde{T}_{1/n}$ and similarly $\mathcal{H}_j(mn^\alpha, mn)$ can be determined by $\gamma_{1/mn}$ upon $\tilde{T}_{1/mn}$. In Lemma 7.38, we will show that the indicator functions of corresponding events of the exploration process are identical under three constraints. Finally, we prove that all of these constraints happen with high probability under the good coupling of $\gamma_{1/n}$, $\gamma_{1/mn}$ and γ . Therefore, the probabilities of $\mathcal{H}_j(n^\alpha, n)$ and $\mathcal{H}_j(mn^\alpha, mn)$ are close.

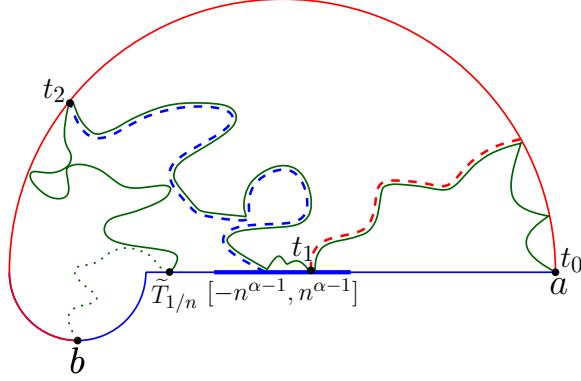


Figure 7.8: Sketch of the event $\mathcal{H}_j(n^\alpha, n)$ when $j = 2$. The whole domain is $\Omega_{[1/n]}$. The exploration process $\gamma_{1/n}$ from $a_{1/n}$ to $b_{1/n}$ on $\Omega_{[1/n]}$ is in green. $\tilde{T}_{1/n}$ is the first time that $\gamma_{1/n}$ hits $[-1, -n^{\alpha-1}]$, after which time $\gamma_{1/n}$ is sketched in dotted. The segment $[-n^{\alpha-1}, n^{\alpha-1}]$ is in bold blue line. Two arms, neighboring the interface $\gamma_{1/n}$, from this segment to the top boundary C_1^+ are in dashed red and blue respectively.

Lemma 7.37. *One has*

$$\mathcal{H}_j(n^\alpha, n) = \{\gamma_{1/n} \text{ travels } j \text{ times between } C_1^+ \text{ and } [-n^{\alpha-1}, n^{\alpha-1}] \text{ before } \tilde{T}_{1/n}\} \quad (7.51)$$

and

$$\mathcal{H}_j(mn^\alpha, mn) = \{\gamma_{1/mn} \text{ travels } j \text{ times between } C_1^+ \text{ and } [-n^{\alpha-1}, n^{\alpha-1}] \text{ before } \tilde{T}_{1/mn}\}. \quad (7.52)$$

By the RHS of (7.51), we mean that there exist $0 = t_0 < t_1 < t_2 < \dots < t_j \leq \tilde{T}_{1/n}$ such that $\gamma_{1/n}(t_0) \in C_1^+$, $\gamma_{1/n}(t_1) \in [-n^{\alpha-1}, n^{\alpha-1}]$, $\gamma_{1/n}(t_2) \in C_1^+$, etc.; the same for (7.52). The proof follows from simple geometrical arguments, so we omit it. See also Figure 7.8 for an illustration. Also, note that if we work with $\mathcal{B}_j(n^\alpha, n)$ instead, we need to deal with a varying domain which is a less-than-ideal strategy. This is the main reason that we introduce and work with $\mathcal{H}_j(n^\alpha, n)$. We refer readers to Remark 7.10 and the paragraph below Proposition 7.4 for more discussions.

Before stating the second lemma, we introduce three events which all happen with high probability under good coupling of $\gamma_{1/n}$, $\gamma_{1/mn}$ and γ . Let $u = u(\Omega)$ be the constant from Proposition 7.35. We will call $C_1^+ \cup [-n^{\alpha-1}, n^{\alpha-1}]$ the **designated boundary** and note that $(-1, 0)$, $(1, 0)$, $(-n^{\alpha-1}, 0)$

and $(n^{\alpha-1}, 0)$ are the extremal points of the designated boundary. Define $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$ as

$$\begin{aligned}\mathcal{J}_1 &= \left\{ d\left(\gamma_{1/n}|_{[0, \tilde{T}_{1/n}]}, \gamma_{1/mn}|_{[0, \tilde{T}_{1/mn}]}\right) \leq 2n^{-u} \right\}; \\ \mathcal{J}_2 &= \left\{ \text{There does not exist } x \text{ on the boundary such that it is } n^{-v} \text{ away from the extremal} \right. \\ &\quad \text{points, and } \gamma_{1/n} \text{ enters } B(x, 2n^{-u}) \text{ but does not hit the designated boundary between} \\ &\quad \text{the last entrance and first exit times of } B(x, n^{-v}), \text{ and the same for } \gamma_{1/mn} \left. \right\}; \\ \mathcal{J}_3 &= \left\{ \gamma_{1/n} \text{ and } \gamma_{1/mn} \text{ do not enter } B((-1, 0), 2n^{-v}), B((-n^{\alpha-1}, 0), 2n^{-v}) \right. \\ &\quad \left. \text{or } B((n^{\alpha-1}, 0), 2n^{-v}) \text{ and do not reenter } B((1, 0), 2n^{-v}) \text{ after leaving } B((1, 0), 1/4) \right\}.\end{aligned}$$

Here, $v \in (0, u)$ is a constant such that (7.55) holds. If \mathcal{J}_2 happens, then each time $\gamma_{1/n}$ (or $\gamma_{1/mn}$) enters $B(x, 2n^{-u})$, it will hit the designated boundary between the last entrance and first exit times of $B(x, n^{-v})$ for all points x on the designated boundary that are n^{-v} away from the extremal points.

We now give the second lemma which states that the indicator functions of $\mathcal{H}_j(n^\alpha, n)$ and $\mathcal{H}_j(mn^\alpha, mn)$ are identical under three constraints. In the proof, we will use the equivalence established in Lemma 7.37.

Lemma 7.38. *One has*

$$\mathbb{1}_{\mathcal{H}_j(n^\alpha, n)} = \mathbb{1}_{\mathcal{H}_j(mn^\alpha, mn)} \text{ on } \mathcal{J}_1 \cap \mathcal{J}_2 \cap \mathcal{J}_3. \quad (7.53)$$

Proof. Assume that $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$ and $\mathcal{H}_j(n^\alpha, n)$ happen so that $\gamma_{1/n}$ fulfills the RHS of (7.51). Suppose that $\gamma_{1/n}$ hits $[-n^{\alpha-1}, n^{\alpha-1}]$ at the point x . From \mathcal{J}_3 we know that x is n^{-v} away from the extremal points. By \mathcal{J}_1 and \mathcal{J}_2 , the process $\gamma_{1/mn}$ enters $B(x, 2n^{-u})$ and so must hit the designated boundary between the last entrance and first exit times of $B(x, n^{-v})$, therefore also before $\tilde{T}_{1/mn}$. Thus, $\gamma_{1/mn}$ must hit $[-n^{\alpha-1}, n^{\alpha-1}]$ at the same place as $\gamma_{1/n}$ before $\tilde{T}_{1/mn}$. Similarly, when $\gamma_{1/n}$ hits a point x in C_1^+ before $\tilde{T}_{1/n}$ and after hitting $[-n^{\alpha-1}, n^{\alpha-1}]$ (this together with \mathcal{J}_3 ensures that the hitting point is $2n^{-v}$ away from the point $(1, 0)$), the process $\gamma_{1/mn}$ must hit C_1^+ at the same place as $\gamma_{1/n}$ before $\tilde{T}_{1/mn}$. Hence, $\gamma_{1/mn}$ fulfilling the RHS of (7.52) and so $\mathcal{H}_j(mn^\alpha, mn)$ happens. Therefore,

$$\mathcal{J}_1 \cap \mathcal{J}_2 \cap \mathcal{J}_3 \cap \mathcal{H}_j(n^\alpha, n) \subset \mathcal{H}_j(mn^\alpha, mn) \text{ and similarly } \mathcal{J}_1 \cap \mathcal{J}_2 \cap \mathcal{J}_3 \cap \mathcal{H}_j(mn^\alpha, mn) \subset \mathcal{H}_j(n^\alpha, n).$$

This completes the proof of (7.53). \square

Proof of Proposition 7.36. First, we give upper bounds to $\mathbb{P}[\mathcal{J}_1^c]$, $\mathbb{P}[\mathcal{J}_2^c]$ and $\mathbb{P}[\mathcal{J}_3^c]$. By Proposition 7.35 and (7.50), under the good coupling of $\gamma_{1/n}$, $\gamma_{1/mn}$ and γ

$$\mathbb{P}[\mathcal{J}_1^c] = \mathbb{P}\left[d\left(\gamma_{1/n}|_{[0, \tilde{T}_{1/n}]}, \gamma_{1/mn}|_{[0, \tilde{T}_{1/mn}]}\right) > 2n^{-u}\right] = O(n^{-u}).$$

We claim that if \mathcal{J}_2^c happens, we have a half-plane 3-arm event from a $(2n^{-u})$ -ball on the designated boundary to distance n^{-v} . Suppose \mathcal{J}_2^c happens for $\gamma_{1/n}$ and some point x on the designated boundary (the case of $\gamma_{1/mn}$ can be handled similarly). We define τ^1 as the last¹⁴ entrance time of $B(x, n^{-v})$, σ^1 as the first hitting time of $B(x, 2n^{-u})$ after τ^1 , σ^2 as the first exit time of $B(x, n^{-v})$ and τ^2 as the last exit time of $B(x, 2n^{-u})$ before σ^2 . Then, the left and right boundaries of $\gamma_{1/n}[\tau^1, \sigma^1]$ and $\gamma_{1/n}[\tau^2, \sigma^2]$ contain three disjoint crossings from $B(x, 2n^{-u})$ to distance n^{-v} . In other words, there are two crossings of the same color at the two boundaries close to $C_1^+ \cup [-n^{\alpha-1}, n^{\alpha-1}]$ and another crossing of a different color in the area sandwiched by the process. Moreover, we can cover the designated boundary by a $(2n^{-u})$ -net with $O(n^u)$ elements. So, for some $c > 0$

$$\mathbb{P}[\mathcal{J}_2^c] \leq O(n^u) \times O(n^{(1+c)(v-u)}). \quad (7.54)$$

If \mathcal{J}_3^c happens, we have a planar 1-arm event from a $(2n^{-v})$ -ball centered at one of the extremal points to distance $1/4$ since we can find an arm along the exploration process that is $2n^{-v}$ close to that extremal point. So,

$$\mathbb{P}[\mathcal{J}_3^c] \leq O(1) \times O(n^{-cv})$$

(where we reduce the value of c if necessary).

Therefore, under the good coupling

$$\begin{aligned} |h_j(n^\alpha, n) - h_j(mn^\alpha, mn)| &\stackrel{(7.53)}{\leq} 1 - \mathbb{P}[\mathcal{J}_1^c \cup \mathcal{J}_2^c \cup \mathcal{J}_3^c] \leq \mathbb{P}(\mathcal{J}_1^c) + \mathbb{P}(\mathcal{J}_2^c) + \mathbb{P}(\mathcal{J}_3^c) \\ &= O(n^{-u}) + O(n^u) \times O(n^{(1+c)(v-u)}) + O(n^{-cv}). \end{aligned}$$

Take a small v such that

$$u + (1+c)(v-u) < 0. \quad (7.55)$$

By applications of the RSW theory, $h_j(mn^\alpha, mn) \geq n^{C(\alpha-1)}$. We can take c_8 such that

$$-Cc_8 > \max\{-u, u + (1+c)(v-u), -cv\}.$$

Then, for all $\alpha \in (1 - c_8, 1)$, $|h_j(n^\alpha, n) - h_j(mn^\alpha, mn)| = O(n^{-c})h_j(mn^\alpha, mn)$ as desired. \square

¹⁴Here, last and first are defined with respect to some hitting time of $B(x, 2n^{-u})$.

Remark 7.39. When $j = 1$, the event $\mathcal{H}_j(r, R)$ is simply equivalent to the event that there is a red crossing between $[-r/R, r/R]$ and C_1^+ in B_1^+ with the lattice scale $\eta = 1/R$. This special case of Proposition 7.36 has been proved in Proposition 5.6 of [18] and the Main Theorem of [2].

7.5.2 Proof of Theorem 7.1

In this subsection, we will prove Proposition 7.4 and complete the proof of Theorem 7.1. We will first prove a version of (7.8) for the case of h_j with Propositions 7.36 and 7.28 as inputs, whose proof is simpler than that of (7.8) for b_j yet already contains the main idea. Then, we will turn back to (7.8) for b_j with Proposition 7.29 as additional inputs.

Proof of Proposition 7.4. As discussed above, we start with the case of h_j . Fix $r \geq r_h(j)$. Let $m \in (1.1, 10)$ and $\alpha \in (1 - c_8, 1)$ (c_8 is the constant defined in Proposition 7.36). In this proof we write $f(n) \simeq g(n)$ as the shorthand of $f(n) = g(n)(1 + O(n^{-c}))$ where the constants may depend on j and on the choice of α but are independent of the choice of m .

We first transform $\frac{h_j(r, mn)}{h_j(r, n)}$ into comparisons of mesoscopic arm probabilities thanks to Proposition 7.28. Since $\mathcal{H}_j(r, n) \subset \mathcal{H}_j(n^\alpha, n)$ and $\mathcal{H}_j(r, mn) \subset \mathcal{H}_j(n^\alpha, mn)$,

$$\frac{h_j(r, mn)}{h_j(r, n)} = \frac{\mathbb{P}[\mathcal{H}_j(r, mn) | \mathcal{H}_j(n^\alpha, mn)] \cdot \mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(r, n) | \mathcal{H}_j(n^\alpha, n)] \cdot \mathbb{P}[\mathcal{H}_j(n^\alpha, n)]} \stackrel{(7.35)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n)]}. \quad (7.56)$$

We then use Proposition 7.36 to pass from scales mn, n to scales m^2n, mn :

$$\frac{\mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n)]} \simeq \frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]}. \quad (7.57)$$

Next, similar to the reverse of the first step, we transform from mesoscopic comparison to the ratio $\frac{h_j(r, m^2n)}{h_j(r, mn)}$ by Proposition 7.28. We have

$$\frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]} \stackrel{(7.35)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]} \cdot \frac{\mathbb{P}[\mathcal{H}_j(r, m^2n) | \mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(r, mn) | \mathcal{H}_j(mn^\alpha, mn)]} = \frac{h_j(r, m^2n)}{h_j(r, mn)}. \quad (7.58)$$

The combination of (7.56), (7.57) and (7.58) gives

$$\frac{h_j(r, mn)}{h_j(r, n)} \stackrel{(7.56)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n)]} \stackrel{(7.57)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]} \stackrel{(7.58)}{\simeq} \frac{h_j(r, m^2n)}{h_j(r, mn)}. \quad (7.59)$$

We now turn back to b_j . First,

$$\frac{b_j(r, mn)}{b_j(r, n)} = \frac{\mathbb{P}[\mathcal{B}_j(r, mn) | \mathcal{B}_j(r, n^\alpha)]}{\mathbb{P}[\mathcal{B}_j(r, n) | \mathcal{B}_j(r, n^\alpha)]} \stackrel{(7.42)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(r, mn) | \mathcal{H}_j(r, n^\alpha)]}{\mathbb{P}[\mathcal{H}_j(r, n) | \mathcal{H}_j(r, n^\alpha)]} \stackrel{(7.56)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n)]}. \quad (7.60)$$

The next step is the same as (7.57). The third step is similar to the reverse of the first step, but we use (7.58) instead of (7.56). As a result, we obtain

$$\frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]} \simeq \frac{b_j(r, m^2n)}{b_j(r, mn)}. \quad (7.61)$$

The combination of the above two estimates on proportions as well as (7.57) gives

$$\frac{b_j(r, mn)}{b_j(r, n)} \stackrel{(7.60)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(n^\alpha, mn)]}{\mathbb{P}[\mathcal{H}_j(n^\alpha, n)]} \stackrel{(7.57)}{\simeq} \frac{\mathbb{P}[\mathcal{H}_j(mn^\alpha, m^2n)]}{\mathbb{P}[\mathcal{H}_j(mn^\alpha, mn)]} \stackrel{(7.61)}{\simeq} \frac{b_j(r, m^2n)}{b_j(r, mn)}. \quad (7.62)$$

This finishes the proof of Proposition 7.4. \square

Proof of Theorem 7.1. For a given j , it suffices to consider only $r \geq r_b(j)$. It follows from Proposition 7.4 and Lemmas 7.13 and 7.14 that there exist $0 < C < \infty$ and $-\infty < \alpha < \infty$ such that $b_j(r, n) = Cn^\alpha(1 + O(n^{-c}))$. By Lemma 7.11, $\alpha = -\beta_j$ and so

$$b_j(r, n) = Cn^{-\beta_j}(1 + O(n^{-c})).$$

And the same holds for h_j . This finishes the proof of Theorem 7.1. \square

7.5.3 Comparison estimates in the plane

Our goal of this subsection is Proposition 7.40 in which we compare the probabilities of \mathcal{Y}_j at different scales. The proof is similar to that of the half-plane case in spirit, in which we relate \mathcal{Y}_j to the exploration process and then make use of the power-law rate of convergence from [20]. However when the number of arms is odd, there are inevitably two neighboring ones of the same color that are not separated by an interface. This issue poses extra difficulties. See Lemma 7.41 and the paragraph right above it for more details.

Fix in this subsection the nice domain (in the sense of Section 7.3.1) and four points on its boundary

$$\Omega = B_1 \cup B\left((0, 1), 2\sin(\pi/12)\right), \quad a = (0, -1), \quad b = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad c = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \text{and } d = (0, 1).$$

See Figure 7.9 for an illustration. We also let

$$e = (0, 1 + 2\sin(\pi/12)) \text{ and } U = B(d, 2\sin(\pi/2)) \setminus B_1 \subset \Omega \text{ which contains a neighborhood of } e.$$

Note that the value $2\sin(\pi/12)$ equals to the length of \overline{bc} (and \overline{cd}). Given a mesh size η , the percolation exploration process starts from a_η and ends at e_η . We assign red color to the counterclockwise

arc \widehat{eba} and blue color to the counterclockwise arc \widehat{ace} . See Figure 7.9 for an illustration of the setup. Recall that $T_{1/n}$ is the first time that $\gamma_{1/n}$ enters U_n . The stopping times $T_{1/mn}$ and T are defined similarly.

Recall (7.15) that $\mathcal{Y}_j(r, R)$ is the event that in C_R with mesh size $\eta = 1$ there exist j disjoint arms connecting C_r with C_R with the prescribed color pattern and some additional constraints. Rescaling the lattice by $1/R$, we get that $\mathcal{Y}_j(r, R)$ is equivalent to the event there exist j such arms connecting $C_{r/R}$ to C_1 in B_1 with mesh size $\eta = 1/R$. Similar to the half-plane case in Section 7.5.1, we will consider two cases $R = n$ and mn in this subsection and we will tacitly assume the equivalence between events on the rescaled and those on the unrescaled lattices.

Recall the definition of $y_j(r, R)$ in (7.16). The following proposition plays a similar role to that of Proposition 7.36 in the half-plane case. It states that y_j are almost the same at different scales within a power-law error when r is not too small compared to R .

Proposition 7.40. *There exists $c_9 > 0$ such that for all $\alpha \in (1 - c_9, 1)$ and $m \in (1.1, 10)$,*

$$y_j(n^\alpha, n) = y_j(mn^\alpha, mn)(1 + O(n^{-c})).$$

Here, $O(n^{-c})$ is independent of m and α .

We postpone the proof till the end of this subsection and turn to Lemmas 7.41 and 7.42 which play same roles as Lemmas 7.37 and 7.38 in the half-plane case. However, the situation here is more complicated, since $\mathcal{Y}_j(n^\alpha, n)$ corresponds to different events according to whether j is even or odd. When j is odd, in the event $\mathcal{Y}_j(n^\alpha, n)$ there are two neighboring arms of the same color which lead to the introduction of the “disjointedness condition” (see (7.64) and (7.66)).

Lemma 7.41. *For even $j \geq 2$, one has*

$$\mathcal{Y}_j(n^\alpha, n) = \{\gamma_{1/n} \text{ hits } C_{n^{\alpha-1}}, \widehat{ba}, C_{n^{\alpha-1}}, \widehat{ac}, \dots, (j-1) \text{ times before } T_{1/n}\}, \quad (7.63)$$

and for odd $j \geq 3$, one has

$$\mathcal{Y}_j(n^\alpha, n) = \left\{ \gamma_{1/n} \text{ hits } C_{n^{\alpha-1}}, \widehat{ba}, C_{n^{\alpha-1}}, \widehat{ac}, \dots, (j-1) \text{ times before } T_{1/n} \text{ and the last two crossings satisfy the “disjointedness condition” defined in (7.66)} \right\}. \quad (7.64)$$

For $\gamma_{1/n}$ satisfying the requirements on the RHS of (7.63) and (7.64), there exists a sequence of times $0 = t_0 < t_1 < t_2 < \dots < t_{j-1} \leq T_{1/n}$ such that

$$\gamma_{1/n}(t_1) \in C_{n^{\alpha-1}}, \quad \gamma_{1/n}(t_2) \in \widehat{ba}, \quad \gamma_{1/n}(t_3) \in C_{n^{\alpha-1}}, \quad \gamma_{1/n}(t_4) \in \widehat{ac}, \quad \text{etc.} \quad (7.65)$$

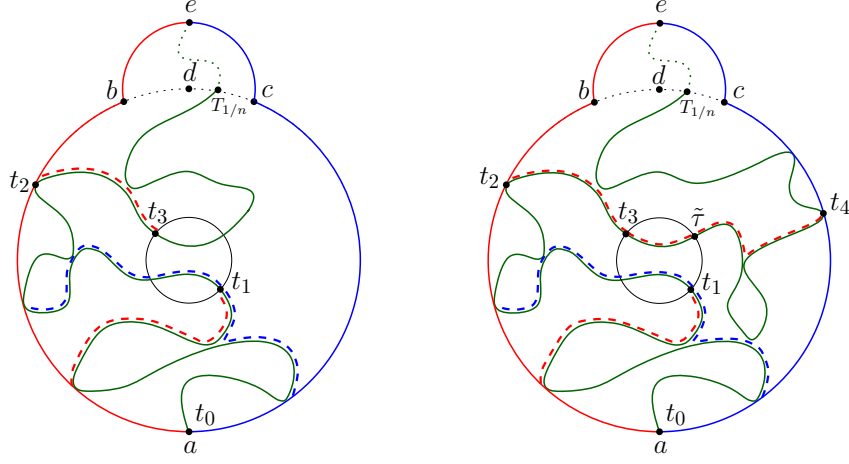


Figure 7.9: Sketch of events $\mathcal{Y}_4(n^\alpha, n)$ on the left and $\mathcal{Y}_5(n^\alpha, n)$ on the right. The whole domain is $\Omega_{[1/n]}$. The small circle inside this domain is $C_{n^{\alpha-1}}$. The exploration process $\gamma_{1/n}$ from $a_{1/n}$ to $e_{1/n}$ on $\Omega_{[1/n]}$ is in green. $T_{1/n}$ is the first time that $\gamma_{1/n}$ enters $U_{[1/n]}$, after which time $\gamma_{1/n}$ is sketched in dotted curves. The arms, neighboring the interface $\gamma_{1/n}$, to enforce the occurrence of arm events are in dashed red and blue. Note that on the right the “disjointedness condition” holds: $\gamma_{1/n}(t_2, t_4)$ does not hit \widehat{cdb} , $\tilde{\sigma} = t_3$, and the left boundaries of $\gamma_{1/n}(t_2, \tilde{\sigma})$ and $\gamma_{1/n}(\tilde{\tau}, t_4)$ are disjoint.

In particular, $\gamma_{1/n}$ is allowed to hit other arcs in the middle of two consecutive hitting times. For instance, it can hit \widehat{ac} between the times hitting $C_{n^{\alpha-1}}$ and \widehat{ba} . We say that this time sequence satisfies the “disjointedness condition” if

$$\gamma_{1/n}(t_{j-3}, t_{j-1}) \text{ does not hit } \widehat{cdb}, \text{ and the left boundaries of } \gamma_{1/n}[t_{j-3}, \tilde{\sigma}] \text{ and } \gamma_{1/n}[\tilde{\tau}, t_{j-1}] \quad (7.66)$$

(which are the last two red arms) are disjoint,

where $\tilde{\sigma}$ denotes the first hitting time of $C_{n^{\alpha-1}}$ after t_{j-3} and $\tilde{\tau}$ denotes the last exit time from $C_{n^{\alpha-1}}$ before t_{j-1} . The proof follows from simple geometrical arguments, so we omit it. One can see the following figure for an illustration.

Similarly, we can show that for even $j \geq 2$,

$$\mathcal{Y}_j(mn^\alpha, mn) = \{\gamma_{1/mn} \text{ hits } C_{n^{\alpha-1}}, \widehat{ba}, C_{n^{\alpha-1}}, \widehat{ac}, \dots, (j-1) \text{ times before } T_{1/mn}\}, \quad (7.67)$$

and the same for odd $j \geq 3$ with the extra “disjointedness condition”.

Before stating the second lemma, we introduce several events $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ (and \mathcal{K}_4 when j is odd) which all happen with high probability under the good coupling from Proposition 7.22 of $\gamma_{1/n}, \gamma_{1/mn}$ and γ , whose joint law¹⁵ we also denote by \mathbb{P} . Let u be the constant from Proposition 7.22. We will call $\widehat{ba} \cup \widehat{ac} \cup \widehat{C}_{n^\alpha-1}$ the **designated boundary** and note that b, a, d are the extremal points of the designated boundary. Define $\mathcal{K}_1, \mathcal{K}_2$ and \mathcal{K}_3 as

$$\begin{aligned}\mathcal{K}_1 &= \left\{ d(\gamma_{1/n}|_{[0, T_{1/n}]}, \gamma_{1/mn}|_{[0, T_{1/mn}]}) > 2n^{-u} \right\}; \\ \mathcal{K}_2 &= \left\{ \text{there does not exist any point } x \text{ on the designated boundary such that it is } n^{-v} \text{ away} \right. \\ &\quad \text{from the extremal points, and } \gamma_{1/n} \text{ enters } B(x, 2n^{-u}) \text{ but does not hit the designated} \\ &\quad \text{hit the designated boundary between the last entrance and first exit times of } B(x, n^{-v}), \\ &\quad \left. \text{and the same for } \gamma_{1/mn} \right\}; \\ \mathcal{K}_3 &= \left\{ \gamma_{1/n} \text{ and } \gamma_{1/mn} \text{ do not enter } B(b, 2n^{-v}), \text{ or } B(d, 2n^{-v}), \text{ and do not reenter } B(a, 2n^{-v}) \right. \\ &\quad \left. \text{after leaving } B(a, 1/4) \right\},\end{aligned}$$

and for odd j 's, define \mathcal{K}_4 as

$$\mathcal{K}_4 = \left\{ \text{there does not exist a time sequence that satisfies (7.65) and the} \right. \\ \left. \text{“disjointedness condition” (7.66) but the last two arms are } (2n^{-u})\text{-close} \right\}.$$

Here, $v \in (0, u)$ is a constant to be chosen later such that (7.71) holds. If \mathcal{K}_2 happens, then each time $\gamma_{1/n}$ (or $\gamma_{1/mn}$) enters $B(x, 2n^{-u})$, it will hit the designated boundary between the last entrance and first exit times of $B(x, n^{-v})$ for all points x on the designated boundary that are n^{-v} away from the extremal points.

We now give the second lemma which states that the indicator functions of $\mathcal{Y}_j(n^\alpha, n)$ and $\mathcal{Y}_j(mn^\alpha, mn)$ are identical under $\cap_i \mathcal{K}_i$. In the proof, we will use the equivalence between arm events and behavior of exploration processes. established in Lemma 7.41 and (7.67).

Lemma 7.42. *One has*

$$\mathbb{1}_{\mathcal{Y}_j(n^\alpha, n)} = \mathbb{1}_{\mathcal{Y}_j(mn^\alpha, mn)} \text{ on } \cap_{i=1}^3 \mathcal{K}_i \text{ (for } j \text{ even) or } \cap_{i=1}^4 \mathcal{K}_i \text{ (for } j \text{ odd)}. \quad (7.68)$$

¹⁵Similar to the half-plane case, here we also consider a probability space in which all three objects are coupled together.

Proof. We begin with the case where j is even. Assume that $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ and $\mathcal{Y}_j(n^\alpha, n)$ happen and so $\gamma_{1/n}$ fulfills the RHS of (7.63). Suppose that $\gamma_{1/n}$ hits $C_{n^{\alpha-1}}$ at the point x . By \mathcal{K}_1 and \mathcal{K}_2 , the process $\gamma_{1/mn}$ enters $B(x, 2n^{-u})$ and so must hit $C_{n^{\alpha-1}}$ between the last entrance and first exit times of $B(x, n^{-v})$. Thus, $\gamma_{1/mn}$ hits $C_{n^{\alpha-1}}$ at the same place. Suppose that $\gamma_{1/n}$ hits \widehat{ba} at the point x . From \mathcal{K}_3 we know that x is n^{-v} away from the extremal points. By \mathcal{K}_1 and \mathcal{K}_2 , the process $\gamma_{1/mn}$ enters $B(x, 2n^{-u})$ and so hits \widehat{ba} between the last entrance and first exit times of $B(x, n^{-v})$, therefore also before $T_{1/mn}$. The same holds for \widehat{ac} . Therefore, $\gamma_{1/mn}$ hits each side at the same place as $\gamma_{1/n}$ before $T_{1/mn}$ and so $\gamma_{1/mn}$ fulfills the RHS of (7.67) which implies that $\mathcal{Y}_j(mn^\alpha, mn)$ happens. Therefore,

$$\mathcal{K}_1 \cap \mathcal{K}_2 \cap \mathcal{K}_3 \cap \mathcal{Y}_j(n^\alpha, n) \subset \mathcal{Y}_j(mn^\alpha, mn).$$

In the same way, we can show that

$$\mathcal{K}_1 \cap \mathcal{K}_2 \cap \mathcal{K}_3 \cap \mathcal{Y}_j(mn^\alpha, mn) \subset \mathcal{Y}_j(n^\alpha, n).$$

This completes the proof of (7.68).

For odd j , we add one more constraint \mathcal{K}_4 to ensure that we can still find a time sequence that satisfies the “disjointedness condition” even after minor perturbation of the process. Then, it is easy to verify that $\mathbb{1}_{\mathcal{Y}_j(n^\alpha, n)} = \mathbb{1}_{\mathcal{Y}_j(mn^\alpha, mn)}$ on $\mathcal{K}_1 \cap \mathcal{K}_2 \cap \mathcal{K}_3 \cap \mathcal{K}_4$, since after a minor perturbation of $2n^{-u}$, the same time sequence (which may differ by a distance of n^{-v}) still satisfies (7.65) and the “disjointedness condition”. \square

We are now ready to prove the main result of this subsection.

Proof of Proposition 7.40. First, we give upper bounds to $\mathbb{P}[\mathcal{K}_1^c]$, $\mathbb{P}[\mathcal{K}_2^c]$, $\mathbb{P}[\mathcal{K}_3^c]$ and $\mathbb{P}[\mathcal{K}_4^c]$. By Proposition 7.22, there exists $u > 0$ such that under the good coupling of $\gamma_{1/n}$, $\gamma_{1/mn}$ and γ

$$\mathbb{P}[\mathcal{K}_1^c] = \mathbb{P}\left[d\left(\gamma_{1/n}|_{[0, T_{1/n}]}, \gamma_{1/mn}|_{[0, T_{1/mn}]}\right) > 2n^{-u}\right] = O(n^{-u}). \quad (7.69)$$

If \mathcal{K}_2^c happens, we have a half-plane 3-arm event from a $(2n^{-u})$ -ball on the designated boundary to distance n^{-v} , see the proof of (7.54) for more details. Thus, $\mathbb{P}[\mathcal{K}_2^c] < O(n^u) \times O(n^{(1+c)(v-u)})$ for some $c > 0$. If \mathcal{K}_3^c happens, we have a planar 1-arm event from a $(2n^{-v})$ -ball centered at one of the extremal points to distance $1/4$. So, $\mathbb{P}[\mathcal{K}_3^c] < O(1) \times O(n^{-cv})$ (where we reduce c if necessary).

Therefore, for even j , under the good coupling

$$\begin{aligned} |y_j(n^\alpha, n) - y_j(mn^\alpha, mn)| &\stackrel{(7.68)}{\leq} \mathbb{P}(\mathcal{K}_1^c) + \mathbb{P}(\mathcal{K}_2^c) + \mathbb{P}(\mathcal{K}_3^c) \\ &= O(n^{-u}) + O(n^u) \times O(n^{(1+c)(v-u)}) + O(n^{-cv}). \end{aligned} \quad (7.70)$$

Now, we give upper bound to $\mathbb{P}[\mathcal{K}_4^c]$. We will call $C_1 \cup C_{n^{\alpha-1}}$ the extended boundary (which is different from the designated boundary we defined before). If \mathcal{K}_4^c happens, we have a whole-plane 6-arm event from a $(2n^{-u})$ -ball in B_1 to distance n^{-v} , or on the extended boundary a half-plane 4-arm event from distance $3n^{-v}$ to $1/4$. We assume that \mathcal{K}_4^c happens for $\gamma_{1/n}$ and $j \equiv 1 \pmod{4}$. The cases for $\gamma_{1/mn}$ or $j \equiv 3 \pmod{4}$ can be treated similarly.

If \mathcal{K}_4^c happens, then the last two red arms, i.e., the left boundaries of $\gamma_{1/n}[t_{j-3}, \tilde{\sigma}]$ and $\gamma_{1/n}[\tilde{\tau}, t_{j-1}]$ are both $(2n^{-u})$ -close to a point x . If x is n^{-v} away from the extended boundary, we write σ^1 (resp. σ^3) for the first hitting time of $C(x, 2n^{-u})$ after t_{j-3} (resp. $\tilde{\tau}$) and τ^1 (resp. τ^3) for the last exit time of $C(x, n^{-v})$ before σ^1 (resp. σ^3). Write σ^2 (resp. σ^4) for the first exit time of $C(x, n^{-v})$ after σ^1 (resp. σ^3) and τ^2 (resp. τ^4) for the last exit time of $C(x, 2n^{-u})$ before σ^2 (resp. σ^4). Then,

$$\tau^1 < \sigma^1 \leq \tau^2 < \sigma^2 < \tau^3 < \sigma^3 \leq \tau^4 < \sigma^4.$$

Furthermore, the left boundaries of $\gamma_{1/n}[\tau^1, \sigma^1], \gamma_{1/n}[\tau^2, \sigma^2], \gamma_{1/n}[\tau^3, \sigma^3], \gamma_{1/n}[\tau^4, \sigma^4]$ are four disjoint red arms from $B(x, 2n^{-u})$ to distance n^{-v} , and the right boundaries of $\gamma_{1/n}[\tau^1, \sigma^1], \gamma_{1/n}[\tau^3, \sigma^3]$ are two disjoint blue arms from $B(x, 2n^{-u})$ to distance n^{-v} . So, there is a whole-plane 6-arm event from a $(2n^{-u})$ -ball in B_1 to distance n^{-v} . If x is n^{-v} -close to a point y on the extended boundary, then the left and right boundaries of $\gamma_{1/n}[\tau^2, \sigma^2], \gamma_{1/n}[\tau^3, \sigma^3]$ (where we replace $2n^{-u}$ in the definition to $2n^{-v}$ and n^{-v} to $1/4$) are four disjoint crossings from $B(y, 3n^{-v})$ to distance $1/4$. We can cover B_1 by a $(2n^{-u})$ -net with $O(n^{2u})$ elements and cover the extended boundary by a $(2n^{-v})$ -net with $O(n^v)$ elements. Thus,

$$\mathbb{P}[\mathcal{K}_4^c] = O(n^{-cv}) + O\left(n^{2u+(2+c)(v-u)}\right), \text{ where we reduce } c \text{ if necessary.}$$

Therefore, for odd $j \geq 3$, under the good coupling

$$\begin{aligned} |y_j(n^\alpha, n) - y_j(mn^\alpha, mn)| &\leq \mathbb{P}[\mathcal{K}_1^c] + \mathbb{P}[\mathcal{K}_2^c] + \mathbb{P}[\mathcal{K}_3^c] + \mathbb{P}[\mathcal{K}_4^c] \\ &\leq O(n^{-u}) + O(n^u) \times O(n^{(1+c)(v-u)}) + O(n^{-cv}) + O\left(n^{2u+(2+c)(v-u)}\right). \end{aligned}$$

Take a small v such that

$$u + (1+c)(v-u) < 0 \quad \text{and} \quad 2u + (2+c)(v-u) < 0. \quad (7.71)$$

By applications of RSW theory, $y_j(mn^\alpha, mn) > n^{C(\alpha-1)}$. We can take c_9 such that

$$-Cc_9 > \max\{-u, -cv, u + (1+c)(v-u), 2u + (2+c)(v-u)\}.$$

Then, for all $\alpha \in (1 - c_9, 1)$, $|y_j(n^\alpha, n) - y_j(mn^\alpha, mn)| = O(n^{-c})y_j(mn^\alpha, mn)$ as desired. \square

7.5.4 Proof of Theorems 7.2 and 7.3

In this subsection, we will complete the proof of Theorems 7.2 and 7.3.

The proof of Proposition 7.5 is the same as that of Proposition 7.4 except that we will replace Proposition 7.28 with Proposition 7.32 and Proposition 7.36 with Proposition 7.40. We are now ready to prove Theorem 7.2.

Proof of Theorem 7.2. We start with y_j . For a given j , it suffices to consider only $r \geq r_y(j)$. Combining with Proposition 7.5 and Lemmas 7.13, 7.14, there exist $0 < C < \infty$ and α such that $y_j(r, n) = Cn^\alpha(1 + O(n^{-c}))$. By Lemma 7.11, $\alpha = -\alpha_j$.

We now turn to x_j . By applying Proposition 7.34 we obtain that

$$\frac{x_j(r, mn)}{x_j(r, n)} = \frac{y_j(r, mn)}{y_j(r, n)} \left(1 + O(n^{-c})\right) \quad (7.72)$$

for r, m satisfying the requirement of Proposition 7.5. We thus establish a proportion estimate similar to (7.9) for x_j , which yields (7.5) for x_j . \square

We now turn to Theorem 7.3.

Proof of Theorem 7.3. By Proposition 7.33,

$$\frac{a_j(1, n)}{x_j(1, n)} \stackrel{(7.46)}{=} \frac{a_j(\epsilon n, n)}{x_j(\epsilon n, n)} \left(1 + O(\epsilon^c)\right). \quad (7.73)$$

Thanks to Claim (4) of Lemma 7.11, the first term of the RHS of (7.73) converges to $g_j(\epsilon)/f_j(\epsilon)$ as $n \rightarrow \infty$. Hence the LHS of (7.73) must also converge to a constant as $n \rightarrow \infty$. Combined with (7.5) for x_j , this finishes the proof. \square

Remark 7.43. Note that although $a_6(1, n) > 0$, we did not include the case $j = 6$ in the statement of Theorem 7.3, because $y_6(1, n) = 0$ by definition and hence (7.72) no longer holds. In this case (plus some other inner initial configurations), obtaining sharp asymptotics for $x_6(1, n)$ as well as for $a_6(1, n)$ requires some extra coupling argument, which we omit for brevity.

7.6 Proof of coupling results

In this section, we give the proof of the couplings in the half-plane and plane cases stated in Propositions 7.27 and 7.30 respectively. As the idea for all these couplings are quite similar, we will give a detailed proof for Proposition 7.27 in Section 7.6.1 but only sketch the differences in the details of the proof for Proposition 7.30 in Section 7.6.2.

Before going into details for each specific setup, we briefly explain here the core idea of these couplings. In essence, one divides the domain into exponential scales (“layers” in the text) and couple configurations of each layer step by step. In each step, when passing from one scale to the next, one can show by the separation lemma and RSW-FKG gluing technique that (although under different types of conditioning) the laws of nice configurations in previous scales are absolutely continuous with respect to each other with bounded Radon-Nykodym derivative, and hence different conditional laws can be coupled with positive probability at each scale. Thus the coupling will succeed in the end with high probability. It can also be proved that with high probability, some good events will happen which encompass all the dependence of the past and future configurations under the conditional laws. The result we want follows from the combination of these two facts.

7.6.1 $j \geq 2$ arms in the half-plane

In this subsection, we give the proof of Proposition 7.27. We divide the (discretized) half-plane \mathbb{H} into the disjoint union of dyadic semi-annuli¹⁶ $A_i = B_{r_{i+1}}^+ \setminus B_{r_i}^+$, where $r_i = 2^i$ for $i = 0, 1, 2, \dots$. Call the A_i ’s the **layers** in \mathbb{H} . We will do the coupling layer by layer.

In the following, we fix some annulus region $A_i \cup A_{i+1}$ between C_r^+ and C_R^+ and consider the percolation configuration inside this annulus. We start by defining the **good event** in $A_i \cup A_{i+1}$.

Definition 7.2 (Good event). *Define the good event $\mathcal{G}_j^{(i)}$ associated with the percolation configuration inside $A_i \cup A_{i+1}$ as follows:*

- *There are exactly $(j-1)$ interfaces crossing the annulus $A_i \cup A_{i+1}$. Denote them by $\gamma_1, \dots, \gamma_{j-1}$ in counterclockwise order, then γ_k is adjacent to γ_{k+1} for any $1 \leq k \leq j-2$. In addition, $\gamma_1, \dots, \gamma_{j-1}$ are well-separated on both ends.*

¹⁶For simplicity in this subsection we do not add the $+$ in the superscript.

- There is a red path connecting $[r_i, 2r_i]$ to γ_1 in A_i and a red path connecting $[2r_i, 4r_i]$ to γ_1 in A_{i+1} .
- There are two red (if j is odd) or blue (if j is even) paths which connect $[-2r_i, -r_i]$ and $[-4r_i, -2r_i]$ to γ_{j-1} in A_i and A_{i+1} , respectively.

We will need that the good event $\mathcal{G}_j^{(i)}$ in $A_i \cup A_{i+1}$ happens with at least some positive probability depending only on j .

Lemma 7.44. *There is a constant $c(j) > 0$ such that for any $r_i \geq 10j$, we have $\mathbb{P}[\mathcal{G}_j^{(i)}] \geq c$.*

This lemma could be proved in a way similar to the second proof of [8, Lemma 3.4] by constructing pivotal points through the exploring process (to ensure the interface conditions) together with FKG-RSW gluing (to ensure the well-separateness conditions). Here we present the following alternative proof which relies on a combination of FKG-RSW gluing and BK-Reimer's inequality.

Proof. Define a **quasi-good** event \mathcal{U} in $A_i \cup A_{i+1}$ as follows:

- There are two configurations of well-separated inner and outer faces, say $\Theta = \{\theta_1, \dots, \theta_j\}$ and $\Theta' = \{\theta'_1, \dots, \theta'_j\}$, which are around $C_{4r_i}^+$ and $C_{r_i}^+$ resp., and lying completely in $\mathbb{H} \setminus B_{3r_i}^+$ and in $B_{2r_i}^+$ resp.
- For each $1 \leq k \leq j$, both θ_k and θ'_k are red (if k is odd) or blue (if k is even). Furthermore, there is a path connecting θ_k, θ'_k together with the aforementioned color.

From FKG-RSW gluing technique we see $\mathbb{P}[\mathcal{U}] \geq c_0$ for some positive constant $c_0(j)$. If \mathcal{U} happens, there will be exactly $(j-1)$ interfaces crossing $A_i \cup A_{i+1}$ which end at the $(2j-2)$ endpoints of Θ, Θ' , respectively. Write

$$\mathcal{W} := \left\{ \text{each interface is adjacent to both of its neighbors} \right\}.$$

By the observation that $\mathcal{U} \cap \mathcal{W} \subset \mathcal{G}_j^{(i)}$ and $\mathcal{U} \cap \mathcal{W}^c \subset \mathcal{U} \square \mathcal{B}_1(2r_i, 3r_i)$ (Recall the notation \square from Section 7.3.2), we conclude that (by applying BK-Reimer's inequality (Lemma 7.8) in the last step)

$$\mathbb{P}[\mathcal{G}_j^{(i)}] \geq \mathbb{P}[\mathcal{U} \cap \mathcal{W}] \geq \mathbb{P}[\mathcal{U}] - \mathbb{P}[\mathcal{U} \square \mathcal{B}_1(2r_i, 3r_i)] \geq \mathbb{P}[\mathcal{U}](1 - \mathbb{P}[\mathcal{B}_1(2r_i, 3r_i)]) \quad (7.74)$$

which is bounded below by some $c(j) > 0$. □

Definition 7.3 (Good set). If $\mathcal{G}_j^{(i)}$ holds, let \mathcal{S} be the union of the following colored hexagons (we keep the notation used in the definition of $\mathcal{G}_j^{(i)}$):

- The hexagons that touch at least one of the $(j - 1)$ interfaces;
- The hexagons in the region enclosed by four curves, namely the innermost red path in A_{i+1} from $[2r_i, 4r_i]$ to γ_1 , the outermost red path in A_i from $[r_i, 2r_i]$ to γ_1 , γ_1 , and the real axis;
- The hexagons in the region enclosed by four curves, namely the innermost red (if j is odd) or blue (if j is even) path in A_{i+1} from $[-4r_i, -2r_i]$ to γ_j , the outermost red (if j is odd) or blue (if j is even) path in A_i from $[-2r_i, -r_i]$ to γ_{j-1} , γ_{j-1} , and the real axis.

If $\mathcal{G}_j^{(i)}$ fails, set $\mathcal{S} = \emptyset$. For a nonempty set S of hexagons in $A_i \cup A_{i+1}$, we say that S is a **good set** if S is a possible value for \mathcal{S} such that $\mathcal{G}_j^{(i)}$ holds. See Figure 7.10 for an illustration.

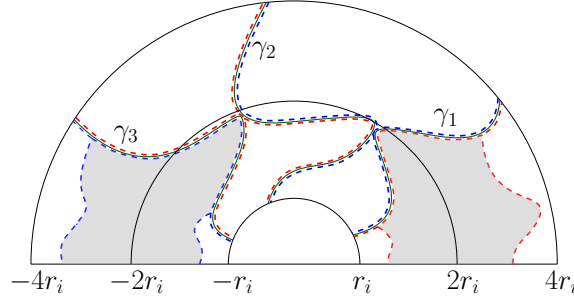


Figure 7.10: The good set \mathcal{S} when $\mathcal{G}_j^{(i)}$ holds with $j = 4$. The three interfaces $\gamma_i, i = 1, 2, 3$ crossing $A^+(r_i, 4r_i)$ are in green. All the hexagons that are adjacent to these interfaces are sketched in dashed red or blue curves, and they make up part of the good set \mathcal{S} . The gray regions on both sides are bounded by the interfaces γ_1 and γ_3 with some innermost paths (crossings) in A_{i+1} and outermost ones in A_i . The (colored) hexagons in these gray regions make up the rest part of \mathcal{S} .

Fix a large integer $K_1(j)$ such that¹⁷

$$p(K_1) := \sup_{i \in \mathbb{N}} \mathbb{P}[\mathcal{B}_{K_1}(r_i, r_{i+1})] < 2^{-10-8\beta_j}. \quad (7.75)$$

¹⁷Note that this is possible since $\beta_j \asymp j^2$; same for the definition of K_2 and K_4 later in this section, noting that $a_j \asymp j^2$ too.

The following proposition is a generalization of (3.1) and (3.2) of [8].

Lemma 7.45. *There are constants $C, C' > 0$ depending only on j , such that for any $2^{K_1+1}r \leq r_i < 4r_i \leq 2^{-K_1-1}R$, the following holds: for any good set $S \subset A_i \cup A_{i+1}$, any configuration of outer faces Θ around $C_{r_i+K_1+3}^+$ with no more than K_1 faces, and any color configuration ω_0 coincides with Θ and satisfies $\mathbb{P}[\mathcal{H}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$,*

$$C^{-1}2^{-|S|} \leq \mathbb{P}[S = S \mid \mathcal{H}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \leq C2^{-|S|}. \quad (7.76)$$

Furthermore,

$$\mathbb{P}[\mathcal{G}_j^{(i)} \mid \mathcal{H}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \geq C'. \quad (7.77)$$

And similar bounds also hold if we replace $\mathcal{H}_j(r, R)$ by $\mathcal{B}_j(r, R)$. In a similar fashion, same results also hold for inner faces Θ around $C_{r_i-K_1-1}^+$ satisfying the same requirements.

Proof. We give the proof of (7.76) and (7.77), and the rest are similar. Rewrite the probability in (7.76) by Bayesian formula as

$$\mathbb{P}[S = S \mid \mathcal{H}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] = 2^{-|S|} \times \frac{\mathbb{P}[\mathcal{H}_j(r, R) \mid S = S, \omega_{\mathcal{D}_\Theta} = \omega_0]}{\mathbb{P}[\mathcal{H}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0]}. \quad (7.78)$$

Let Θ_1 (resp. Θ_2) be the configuration of inner faces around $C_{4r_i}^+$ (resp. the configuration of outer faces around $C_{r_i}^+$) induced by S . Denote \mathcal{R} for the event that there are j arms connecting C_R^+ to each face of Θ_1 , and \mathcal{S} for the event that there are j arms connecting $[-r, r]$ to each face of Θ_2 (each arm has the same color with the face it connects to). Then we see that the event in numerator of (7.78) is equivalent to $\mathcal{R} \cap \mathcal{S}$ conditioned on $\{\omega_{\mathcal{D}_{\Theta_1}} = \omega_0\}$. By independence it has probability $\mathbb{P}[\mathcal{R} \mid \omega_{\mathcal{D}_{\Theta_1}} = \omega_0] \cdot \mathbb{P}[\mathcal{S}]$.

Since Θ_2 is well-separated, by the separation lemma (Lemma 7.15) and FKG-RSW gluing,

$$\mathbb{P}[\mathcal{S}] \asymp \mathbb{P}[\mathcal{H}_j(r, r_{i-1})]. \quad (7.79)$$

In addition, by (7.23) in Proposition 7.17 and FKG-RSW gluing, we also have

$$\mathbb{P}[\mathcal{R} \mid \omega_{\mathcal{D}_{\Theta_1}} = \omega_0] \asymp \mathbb{P}[\mathcal{B}_j(r_{i+3}, R) \mid \omega_{\mathcal{D}_{\Theta_1}} = \omega_0]. \quad (7.80)$$

Finally, similar to the proof of Proposition 7.17, we can show the following quasi-multiplicativity

$$\mathbb{P}[\mathcal{B}_j(r_{i+3}, R) \mid \omega_{\mathcal{D}_{\Theta_1}} = \omega_0] \cdot \mathbb{P}[\mathcal{H}_j(r, r_{i-1})] \asymp \mathbb{P}[\mathcal{H}_j(r, R) \mid \omega_{\mathcal{D}_{\Theta_1}} = \omega_0]. \quad (7.81)$$

Combining (7.78) and (7.79), (7.80) and (7.81) we see that (7.76). Summing over all good set $S \subset A_i \cup A_{i+1}$, and using Lemma 7.44, we get (7.77) as desired. \square

Proof of Proposition 7.27. We focus on the first claim in Proposition 7.27. Recall that $u = \sqrt{rR}$. Let $i_0 < i_N$ be the integers such that $r_{i_0-1} < u \leq r_{i_0} < r_{i_N} \leq R < r_{i_N+1}$. We inductively couple the conditional laws $\mathbb{P}_1 = \mathbb{P}[\cdot | \mathcal{H}_j(r, R)]$ and $\mathbb{P}_2 = \mathbb{P}[\cdot | \mathcal{H}_j(r, mR)]$ from outside to inside layer by layer. The virtue of the coupling is same as the maximal coupling for a Markov chain on a finite state space.

For initialization, sample the color configuration ω_1, ω_2 outside $C_{i_N}^+$ according to \mathbb{P}_1 and \mathbb{P}_2 independently, and set an index $I = 0$. Assume now we have sampled two color configurations ω_1 and ω_2 outside $C_{r_{i+K_1+1}}^+$ for some $i \geq i_0$ according to some coupling of these two conditional distribution, we proceed as follow:

If $I \geq 0$, then for both configurations, we independently explore all the interfaces crossing $A^+(r_{i+K_1}, r_{i+K_1+1})$ from outside to inside and stop exploring when reaching $\mathbb{R} \cup C_{r_{i+K_1}}^+$. This exploring process would induce two configurations of outer faces around $C_{r_{i+K_1}}^+$ for both $\omega_i, i = 1, 2$, denoted by $\hat{\Theta}_i, i = 1, 2$. Also note that $\mathcal{V}_{\hat{\Theta}_1}$ and $\mathcal{V}_{\hat{\Theta}_2}$ are left unexplored. If either $\hat{\Theta}_1$ or $\hat{\Theta}_2$ has more than K_1 faces, we just keep I unchanged, explore all the hexagons outside $C_{r_{i+K_1}}^+$ and proceed the previous procedure on the scale $C_{r_{i+K_1}}$. Otherwise, we add I by 1, and by applying Lemma 7.45 for both $\mathcal{H}_j(r, R)$ and $\mathcal{H}_j(r, mR)$, we can construct a coupling \mathbb{Q}_i of the laws $\mathbb{P}_1[\cdot | \omega_{\mathcal{D}_{\hat{\Theta}_1}} = \omega_1]$ and $\mathbb{P}_2[\cdot | \omega_{\mathcal{D}_{\hat{\Theta}_2}} = \omega_2]$ with the following property:

$$\mathbb{Q}_i[\mathcal{S}(\omega_1) = \mathcal{S}(\omega_2) \neq \emptyset] > C'', \text{ For some } C''(j) > 0, \quad (7.82)$$

where $\mathcal{S}(\omega_1), \mathcal{S}(\omega_2)$ are the sets induced by ω_1, ω_2 in $A_i \cup A_{i+1}$, respectively.

We sample the pair of color configurations (ω_1, ω_2) outside $C_{r_i}^+$ from \mathbb{Q}_i , and perform the following exploring process to detect whether $\{\mathcal{S}(\omega_1) = \mathcal{S}(\omega_2) \neq \emptyset\}$ holds:

- Step 1** For each of the endpoints x of $\hat{\Theta}_1$, we explore the interface of ω_1 inside $B_{r_{i+K_1}}^+$ starting from x , and stop exploring until it hits $\mathbb{R} \cup C_{r_i}^+$. Denote Γ for the subset of the aforementioned interfaces reaching $C_{r_i}^+$. Then we can check whether $|\Gamma| = j - 1$ and each interface in Γ is adjacent to its neighbors in $A^+(r_i, 4r_i)$.
- Step 2** If the condition in **Step 1** holds, denote γ_1 and γ_{j-1} for the leftmost and rightmost interface in Γ . In the quad Q_1 enclosed by $C_{4r_i}^+, \gamma_1, C_{r_i}^+$ and \mathbb{R} , we start from $Q_1 \cap C_{2r_i}^+$ and explore in ω_1 to find the innermost red path in A_{i+1} connecting $[2r_i, 4r_i]$ to γ_1 , and the outermost red path in A_i connecting $[r_i, 2r_i]$ to γ_1 . Our exploring process stops whenever we find such paths, or if we cannot find them, we stop exploring until reaching $C_{r_i}^+$ or $C_{4r_i}^+$. Perform the same

exploration on the left side of $A_i \cup A_{i+1}$. Up to now, we already have enough information to fix $\mathcal{S}(\omega_1)$ and to assert whether the good event $\mathcal{G}_j^{(i)}$ happens or not for ω_1 .

Step 3 Run the same exploring process for ω_2 . Now we can check whether $\mathcal{G}_j^{(i)}$ happens for both ω_1 and ω_2 and $\mathcal{S}(\omega_1) = \mathcal{S}(\omega_2)$. If so, we set $I = -1$; otherwise we keep I unchanged.

If I is still non-negative, we explore the entire color configuration ω_1, ω_2 outside $C_{r_i}^+$, and proceed the procedure above on the scale r_i . Otherwise $I = -1$, then the identical nonempty good set $\mathcal{S}(\omega_1) = \mathcal{S}(\omega_2)$ induces a configuration of outer faces Θ^* with j faces around $C_{r_i}^+$, which is common for both $\omega_i, i = 1, 2$. Note that our exploring process ensures that none of the hexagons in \mathcal{V}_{Θ^*} are explored, i.e. the law in \mathcal{V}_{Θ^*} is still the critical Bernoulli percolation, and thus by the domain Markov property, $\mathbb{P}[\cdot \mid \mathcal{H}_j(r, R), \omega_{\mathcal{D}_{\Theta^*}} = \omega_1]$ and $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR), \omega_{\mathcal{D}_{\Theta^*}} = \omega_2]$ are precisely equal. Hence, we can couple the color configuration in \mathcal{V}_{Θ^*} of these two distributions identically. In this case we have (ω_1, ω_2) satisfies the requirement in proposition 7.27, and Θ^* is a stopping set as desired.

Finally, we control the probability that our coupling remains unsuccessful until scale r_{i_0} . Write $N_1 = \frac{N-2K_1}{2K_1+2}$, then $N_1 \geq \frac{1}{4K_1+4} \log_2 \frac{R}{r} - 2$ since $N \geq \frac{1}{2} \log_2 \frac{R}{r} - 1$. On the one hand, note that in each time I is plus by 1, we always have a positive probability C'' to couple successfully, no matter how things went in previous layers. Thus

$$\mathbb{P}[I \geq N_1] \leq (1 - C'')^{N_1} \leq (r/R)^\delta \text{ for some } \delta(j) > 0.$$

On the other hand, $I < N_1$ implies there are more than $N - (K_1 + 1)N_1 - K_1 = N/2$ integers $i \in [i_0, i_N - 1]$ satisfying $\mathbf{1}_{\mathcal{B}_{K_1}(r_i, r_{i+1})}(\omega_1) + \mathbf{1}_{\mathcal{B}_{K_1}(r_i, r_{i+1})}(\omega_2) \geq 1$, so

$$\mathcal{T} := \left\{ \sum_{i=i_0}^{i_N-1} \mathbf{1}_{\mathcal{B}_{K_1}(r_i, r_{i+1})} \geq N/4 \right\}$$

happens for either both $\omega_i, i = 1, 2$. Combining

$$\mathbb{P}_1[\mathcal{T}] \leq 2^N p(K_1)^{N/4} / \mathbb{P}[\mathcal{H}_j(r, R)], \quad \mathbb{P}_2[\mathcal{T}] \leq 2^N p(K_1)^{N/4} / \mathbb{P}[\mathcal{H}_j(r, mR)],$$

with the a priori estimates of arm probabilities, we obtain

$$\mathbb{P}[0 \leq I \leq N_1] \leq 2^N p(K_1)^{N/4} \left[(R/r)^{-\beta_j + o(1)} + (mR/r)^{-\beta_j + o(1)} \right],$$

which is bounded by $(r/R)^\delta$ for some $\delta(j) > 0$ from the choice of K_1 in (7.75). Altogether we conclude $\mathbb{P}[I \geq 0] \leq (r/R)^\delta$ for some $\delta(j) > 0$, as desired.

With these tools in hand, we can rerun the coupling from inside to outside to conclude the proof of the second claim of Proposition 7.27. \square

7.6.2 Coupling arm events in the plane

In this subsection, we prove Proposition 7.30. Recall the definitions of events $\mathcal{X}_j(r, R)$, $\mathcal{Y}_j(r, R)$ and $\mathcal{A}_j(r, R)$. We shall couple $\mathbb{P}[\cdot \mid \mathcal{X}_j(r, R)]$ with $\mathbb{P}[\cdot \mid \mathcal{A}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, R)]$ with $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR)]$. As discussed at the beginning of this section, we will just sketch the proofs and note necessary modifications of arguments from Section 7.6.1. For clearance, the cases of odd j 's and even j 's are dealt with separately.

As before, we divide the plane into the union of dyadic annuli $A_i = B_{r_{i+1}} \setminus B_{r_i}$, where $r_i = 2^i$ for $i = 0, 1, 2, \dots$. Call them the **layers** in the plane.

Inward coupling: the odd j case We begin with the case for odd j . As mentioned in Remark 7.31, the major difference between the plane and the half-plane cases lies in (7.44), which is not obvious from the existence of a single configuration of face Θ^* . But as we have shown in the proof of Proposition 7.27, we can indeed construct a coupling \mathbb{Q} such that ω_1, ω_2 are identical on some “good set” with high probability under \mathbb{Q} . In this subsection, we define the plane version of good events and good sets, and obtain similar result as before; we shall see the existence of a common good set is enough to give (7.44).

Now for each $r_i > 10j$, define good event in $A_i \cup A_{i+1}$ as follows:

- There are exactly $(j-1)$ interfaces crossing the annulus, say $\gamma_1, \dots, \gamma_{j-1}$, in counterclockwise order and they are well-separated on both sides of the annulus.
- for $1 \leq k \leq j-2$, γ_k is adjacent to γ_{k+1} , while γ_1 is *not* adjacent to γ_{j-1} .
- The hexagons between γ_1 and γ_{j-1} that touch $\gamma_1 \cup \gamma_{j-1}$ are all red (if $j \equiv 1 \pmod{4}$) or blue (if $j \equiv 3 \pmod{4}$). Further, there are two paths with the aforementioned color connecting γ_1 and γ_{j-1} , which lie in A_i and A_{i+1} , respectively.

And if $\mathcal{G}_j^{(i)}$ holds, let \mathcal{S} be the union of all hexagons which touch at least one of the interfaces $\gamma_k, 1 \leq k \leq j-1$, or lie in the quad enclosed by γ_1, γ_{j-1} , the outermost and innermost paths connecting γ_1 and γ_{j-1} in A_i and A_{i+1} , respectively. We set $\mathcal{S} = \emptyset$ if $\mathcal{G}_j^{(i)}$ fails, and for a nonempty

set of colored hexagons in $A_i \cup A_{i+1}$, we say that S is a **good set** if S is a possible value of \mathcal{S} . We have the following estimates similarly as before:

- There exists $c(j) > 0$, such that for any $r_i > 10j$, we have $\mathbb{P}[\mathcal{G}_j^{(i)}] > c$.
- For any fixed large integer K_2 , there exists $C = C(j, K_2) > 0$, such that for any good set S in $A_i \cup A_{i+1}$ with $2^{K_2+1}r \leq r_i < r_{i+2} \leq 2^{-K_2-1}R$, any configuration of outer faces Θ around $C_{r_i+K_2+3}^+$ with no more than K_2 faces and any color configuration ω_0 which satisfies $\mathbb{P}[\mathcal{A}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$,

$$C^{-1}2^{-|S|} \leq \mathbb{P}[\mathcal{S} = S \mid \mathcal{A}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \leq C2^{-|S|}. \quad (7.83)$$

As a result, $\mathbb{P}[\mathcal{G}_j^{(i)} \mid \mathcal{A}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \geq cC^{-1}$.

- In addition, similar results also hold for $\mathcal{A}_j(r, R)$ replaced by $\mathcal{X}_j(r, R)$ or $\mathcal{Y}_j(r, R)$.

Sketch of Proof for Proposition 7.30, odd j case. Fix a large integer $K_2 = K_2(j)$ such that

$$p(K_2) := \sup_{i \in \mathbb{N}} \mathbb{P}[\mathcal{A}_{K_2}(r_i, r_{i+1})] < 2^{-10-8\alpha_j}.$$

The first two items can be established similarly as in Lemma 7.44 and 7.45, results for $\mathcal{X}_j(r, R)$ and $\mathcal{Y}_j(r, R)$ can also be obtained from the same manner. From these estimates, following the framework of the proof of Proposition 7.27, we can construct a coupling of $\mathbb{P}[\cdot \mid \mathcal{A}_j(r, R)]$ (resp. $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, R)]$) and $\mathbb{P}[\cdot \mid \mathcal{X}_j(r, R)]$ (resp. $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR)]$) such that if we sample (ω_1, ω_2) according to such coupling, then with probability at least $(1 - (r/R)^\delta)$ for some $\delta(j) > 0$, ω_1 and ω_2 are identical on some good set outside $C_{\sqrt{rR}}$. This good set induces a common configuration of faces in both ω_i , $i = 1, 2$, which is a stopping set and it also ensures (7.43) (resp. (7.44)), thus completing the proof. \square

Inward coupling: even j case Finally, we deal with the case for even j . This is essentially the case studied in [8]; we follow the idea in this paper and define good events $\mathcal{G}_j^{(i)}$ in $A_i \cup A_{i+1}$ as

- There are exactly j interfaces crossing the annulus $A_i \cup A_{i+1}$, and they are well-separated on both ends. In addition, each of the interfaces is adjacent with its two neighbors.

If $\mathcal{G}_j^{(i)}$ holds, let \mathcal{S} be the union of all colored hexagons in $A_i \cup A_{i+1}$ which touch at least one of the j interfaces, otherwise we set $\mathcal{S} = \emptyset$. For a nonempty set S of colored hexagons in $A_i \cup A_{i+1}$, we say that S is a **good set** if S is a possible value of \mathcal{S} .

The proof of the first item in Proposition 7.30 is essentially same as before, since coinciding on some good set implies (7.43). However, for the second item, in order to guarantee (7.44), it is not enough to make ω_1 and ω_2 coincide on some good set. Indeed, we need to further specify the connecting pattern between this good set and the outer boundary C_R (or C_{mR}). For a color configuration ω_0 sampled from $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, R)]$ (resp. $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR)]$), assume that under ω_0 , for some $u \in [r, R]$ there is a configuration of outer faces Θ around C_u with j faces. Denote $\tilde{\theta}$ for the first red face in Θ when counting from $(0, -u)$ counterclockwise, and denote $\hat{\theta}$ for the red face in Θ which is connected to C_R (resp. C_{mR}) by the first red arm counting from $(0, -R)$ (resp. $(0, -mR)$) counterclockwise. Then we have the following estimates:

- There exists $c(j) > 0$, such that for any $r_i > 10j$, we have $\mathbb{P}[\mathcal{G}_j^{(i)}] > c$.
- Fix a large integer K_3 , there exists $C = C(j) > 0$, such that for any good set S in $A_i \cup A_{i+1}$ with $2^{K_3+1}r \leq r_i < r_{i+2} < 2^{-K_3-1}R$, any configuration of outer faces Θ around $C_{r_i+K_3+3}^+$ with no more than K_3 faces and any color configuration ω_0 satisfies $\mathbb{P}[\mathcal{Y}_j(r, R) \mid \omega_{\mathcal{D}_\Theta} = \omega_0] > 0$, denoting Θ_S for the configuration of outer faces around C_{r_i} induced by S , then it holds that

$$C^{-1}2^{-|S|} \leq \mathbb{P}[S = S, \tilde{\theta}_S = \hat{\theta}_S \mid \mathcal{Y}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \leq C2^{-|S|}. \quad (7.84)$$

As a result, $\mathbb{P}[\mathcal{G}_j^{(i)}, \tilde{\theta}_S = \hat{\theta}_S \mid \mathcal{Y}_j(r, R), \omega_{\mathcal{D}_\Theta} = \omega_0] \geq cC^{-1}$.

Sketch of Proof for Proposition 7.30, even j case. Fix a large integer $K_3(j)$ such that

$$p(K_3) := \sup_{i \in \mathbb{N}} \mathbb{P}[\mathcal{A}_{K_3}(r_i, r_{i+1})] \leq 2^{-10-8\alpha_j}.$$

The estimates above can be proved similarly, with the caveat that for (7.84), we construct some specific gluing to make sure that $\tilde{\theta}_S$ coincides with $\hat{\theta}_S$.

Note that if two color configurations ω_1 and ω_2 are sampled from $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, R)]$ and $\mathbb{P}[\cdot \mid \mathcal{Y}_j(r, mR)]$ respectively, such that $\omega_i, i = 1, 2$ share a common good set S , and $\tilde{\theta}_S = \hat{\theta}_S$ holds for both configurations, then we can conclude that (7.44) is true and $\Theta^* = \Theta_S$ is also a stopping set. Following the same framework, this observation together with the estimates above gives us the desired result. \square

7.7 Complimentary proofs

7.7.1 Proof of Lemmas 7.11 and 7.13

Proof of Lemma 7.11. We begin with the proof of $h_j(n) \asymp b_j(n)$. Since $\mathcal{H}_j(r, n) \subset \mathcal{B}_j(r, n)$, we have $h_j(n) \leq b_j(n)$. It suffices to prove that $h_j(n) \geq cb_j(n)$ for some constant $c > 0$ and all n large enough. Let $r' = 10j \vee r$ and Γ be the set of interfaces connecting C_n^+ and $C_{r'}^+$ in $A^+(r', n)$. Then,

$$\begin{aligned} h_j(r, n) &\geq \mathbb{P}[\mathcal{H}_j(r, n) | Q_{\text{in}}(\Gamma) > j^{-1}, \mathcal{B}_j(r', n)] \times \mathbb{P}[Q_{\text{in}}(\Gamma) > j^{-1} | \mathcal{B}_j(r', n)] \times \mathbb{P}[\mathcal{B}_j(r', n)] \\ &\geq c \times c \times b_j(r', n) \geq cb_j(r, n). \end{aligned} \quad (7.85)$$

(The second inequality follows from RSW theory and Lemma 7.15. The third inequality follows from the fact that $\mathcal{B}_j(r, n) \subset \mathcal{B}_j(r', n)$.) Claim (2) is derived in the same way, but using Proposition 7.20 instead of Lemma 7.15.

Now, we consider Claim (3). We can show the up-to-constants equivalence between arm probabilities using similar arguments in (7.85). The methods in [25] for calculating asymptotics of arm probabilities down to microscopic scales (i.e., $b_j(n)$ or $p_j(n)$ in our notation) can also be applied to derive mesoscopic asymptotics.

Claim (4) follows from [24] and several stability arguments (which need looser estimates than those in the article). One can refer to the proof of Lemma 2.9 in [8] for more details. \square

Proof of Lemma 7.13. We only prove the case of h_j . The cases for other arm probabilities can be proved in the same way. Let Γ be the set of interfaces connecting $[-r, r]$ and C_n^+ in B_n^+ . For $\epsilon > 0$, define $\mathcal{H}_j^\epsilon(r, n)$ by

$$\mathcal{H}_j^\epsilon(r, n) = \mathcal{H}_j(r, n) \cap \{Q_{\text{ex}}(\Gamma) \geq \epsilon\}.$$

Let \mathcal{R}^ϵ denote the event that there exists a point x on C_n^+ such that there are three disjoint arms (not all of the same color) connecting $C(x, \epsilon n)$ and $C(x, n/2)$. By Claim (3) in Lemma 7.11, $\mathbb{P}[\mathcal{R}^\epsilon] \leq O(1/\epsilon) \times O(\epsilon^{1+c}) = O(\epsilon^c)$. Thanks to the spatial independence of percolation

$$\mathbb{P}[\mathcal{H}_j(r, n) \setminus \mathcal{H}_j^\epsilon(r, n)] \leq \mathbb{P}[\mathcal{H}_j(r, n/2)] \cdot \mathbb{P}[\mathcal{R}^\epsilon] \leq C\mathbb{P}[\mathcal{H}_j(r, n)] \times O(\epsilon^c) = O(\epsilon^c)\mathbb{P}[\mathcal{H}_j(r, n)].$$

Therefore,

$$\mathbb{P}[\mathcal{H}_j^\epsilon(r, n) | \mathcal{H}_j(r, n)] \geq 1 - O(\epsilon^c). \quad (7.86)$$

For all $n \leq t \leq (1 + \epsilon/K)n$, $\mathbb{P}[\mathcal{H}_j(r, t) | \mathcal{H}_j^\epsilon(r, n)] \geq 1 - j(1 - c)^K$. This is because on the event $\mathcal{H}_j^\epsilon(r, n)$ each outer face has a length of at least ϵn which can be partitioned into at least K pieces with a

length of at least $(1+\epsilon/K)n$, and by applications of RSW theory, each piece has a probability at least c to be connected with $C_{(1+\epsilon/K)n}^+$, independently of others. Therefore, for all $n \leq t \leq (1+\epsilon/K)n$,

$$\begin{aligned} h_j(t) &= \mathbb{P}[\mathcal{H}_j(r, t)] \geq \mathbb{P}[\mathcal{H}_j(r, t) | \mathcal{H}_j^\epsilon(r, n)] \times \mathbb{P}[\mathcal{H}_j^\epsilon(r, n) | \mathcal{H}_j(r, n)] \times \mathbb{P}[\mathcal{H}_j(r, n)] \\ &\geq (1 - j(1 - c)^K)(1 - O(\epsilon^c))h_j(r, n). \end{aligned}$$

Picking K large and ϵ small, we then complete the proof. \square

7.7.2 Proof of Lemma 7.14

Proof of Lemma 7.14. For $m \in (1.1, 10)$ and integer $k \geq 1$, let $b_k(m) = a_{m^{k+1}}/a_{m^k}$. We note that all constants in $O(\cdot)$'s in this proof are uniform w.r.t. $m \in (1.1, 10)$. By Assumption (1) in the statement, $b_{k+1}(m) = b_k(m)(1 + O(m^{-ck}))$. Then, it is easy to see that $\lim_{k \rightarrow \infty} b_k(m)$ exists (which will be denoted as $C(m)$) and is positive (since $b_k(m)$ is positive when k is sufficiently large). In addition,

$$b_k(m) = C(m) \left(1 + O(m^{-ck})\right), \quad \text{and equivalently} \quad \frac{a_{m^{k+1}}}{C(m)^{k+1}} = \frac{a_{m^k}}{C(m)^k} \left(1 + O(m^{-ck})\right).$$

Therefore, $\lim_{k \rightarrow \infty} \frac{a_{m^k}}{C(m)^k}$ exists (which will be denoted by $B(m)$) and is positive. In addition,

$$a_{m^k} = B(m)C(m)^k \left(1 + O(m^{-ck})\right). \quad (7.87)$$

Next, we use Assumption (2) to prove that there exist $\alpha \in (-\infty, \infty)$ and $C \in (0, \infty)$ such that $C(m) = m^\alpha$ and $B(m) = C$ for all m . It suffices to show that

$$\log(C(m_1))/\log(m_1) = \log(C(m_2))/\log(m_2) \text{ and } B(m_1) = B(m_2) \text{ for all } m_1, m_2 \in (1.1, 10).$$

Fix $m_1, m_2 \in (1.1, 10)$. WLOG we can assume that $\log_{m_2}(m_1)$ is irrational, because we can always find another m_3 such that $\log_{m_3}(m_1)$ and $\log_{m_3}(m_2)$ are both irrational and the case of m_1, m_2 follows from those of m_1, m_3 and m_2, m_3 . Let $\delta, \epsilon > 0$ denote two small constants to be chosen later. Since $\log_{m_2}(m_1)$ is irrational, we can find a sequence of increasing integers $\{p_j\}_{j \geq 1}$ and $\{q_j = \lfloor p_j \log_{m_2}(m_1) \rfloor\}_{j \geq 1}$ such that for all $j \geq 1$

$$q_j \leq p_j \log_{m_2}(m_1) \leq q_j + \epsilon. \quad (7.88)$$

By Assumption (2), there exists a constant $c = c(\delta) > 0$ such that for all $\epsilon < c$

$$\liminf_{n \rightarrow \infty} \inf_{10^{-\epsilon}n \leq s \leq t \leq n} \frac{a_t}{a_s} > 1 - \delta.$$

By (7.88) and the fact that $m_2 < 10$, we have $10^{-\epsilon} m_1^{p_j} \leq m_2^{-\epsilon} m_1^{p_j} \leq m_2^{q_j} \leq m_1^{p_j}$ and so for j large enough

$$(1 - \delta) a_{m_1^{p_j}} \leq a_{m_2^{q_j}} \leq \frac{1}{1 - \delta} a_{m_1^{p_j}}.$$

Together with (7.87)

$$(1 - \delta) \left(1 + O(m_1^{-cp_j}) + O(m_2^{-cq_j}) \right) \leq \frac{B(m_2)C(m_2)^{q_j}}{B(m_1)C(m_1)^{p_j}} \leq \frac{1}{1 - \delta} \left(1 + O(m_1^{-cp_j}) + O(m_2^{-cq_j}) \right).$$

Let j tend to infinity. Since $\lim_{j \rightarrow \infty} \frac{q_j}{p_j} = \log_{m_2}(m_1)$, we have

$$\alpha := \log(C(m_1)) / \log(m_1) = \log(C(m_2)) / \log(m_2).$$

Note that when $\alpha \geq 0$, $10^{-\alpha\epsilon} m_1^{\alpha p_j} \leq m_2^{\alpha q_j} \leq m_1^{\alpha p_j}$; when $\alpha < 0$, $10^{-\alpha\epsilon} m_1^{\alpha p_j} \geq m_2^{\alpha q_j} \geq m_1^{\alpha p_j}$. Therefore,

$$(1 - \delta) 10^{-\epsilon(\alpha \wedge 0)} B(m_1) \left(1 + O(m_1^{-cp_j}) \right) \leq B(m_2) \leq \frac{1}{1 - \delta} B(m_1) 10^{\epsilon(\alpha \vee 0)} \left(1 + O(m_1^{-cp_j}) \right).$$

Let j tend to infinity, then ϵ to zero and finally δ to zero. We have $B(m_1) = B(m_2)$. This completes the proof. \square

7.7.3 Proof of the strong separation lemma

Proof of Proposition 7.16. The proof we employ here is a combination of techniques used in [23] and [8]. We only prove the first bullet point (7.21). Write $s = \text{dist}(r, x_2, \dots, x_j, -r)$. Then $Q(\Theta) = s/r$. Let $M = \lfloor \log_2(r/s) \rfloor \vee 0$ and $L = \lfloor \log_2(R/r) \rfloor$. Set

$$r_i = \begin{cases} r, & i = 0; \\ r + 2^{i-1}s, & 1 \leq i \leq M; \\ 2^{i-M}r, & M + 1 \leq i \leq M + L. \end{cases}$$

Let Γ_i be the collection of interfaces in Γ truncated at their first hitting on $C_{r_i}^+$. Recall that $Q_{\text{ex}}(\Gamma_i)$ is the exterior quality of Γ_i on $C_{r_i}^+$ defined in (7.18). Let $d_i := r_i Q_{\text{ex}}(\Gamma_i)$ be the minimal distance between $-r_i$, r_i and the endpoints of Γ_i . Define the **relative qualities**

$$Q^*(i) := \begin{cases} d_i / (2^i s), & 0 \leq i \leq M; \\ Q_{\text{ex}}(\Gamma_i), & M + 1 \leq i \leq M + L. \end{cases}$$

We set $Q^*(i) := 0$ if not all j interfaces manage to reach $C_{r_i}^+$. Furthermore, set

$$f_i := \mathbb{P}[Q^*(i) > 0], \quad g_i(\rho) := \mathbb{P}[Q^*(i) > \rho] \text{ for } \rho > 0.$$

With the above definitions, it suffices to show

$$g_{M+L}(j^{-1}) \geq c(j)f_{M+L} \tag{7.89}$$

for some $c(j) > 0$ that only depends on j . To this end, we need the following facts about the quality, which can be obtained by using RSW-FKG gluing techniques, similar to the appendices of [23, 8] for the plane case. We omit the details.

- By application of the RSW theory, there is $c_{10}(j) > 0$ such that

$$f_1 \geq c_{10}. \tag{7.90}$$

- For any $\delta > 0$, there exists $\rho_0(\delta, j)$ such that for all $\rho \leq \rho_0$ and $i \geq 0$,

$$f_{i+1} - g_{i+1}(\rho) \leq \delta f_i. \tag{7.91}$$

- For any $\rho > 0$, there exists $R(\rho, j) > 0$ such that for all $i \geq 0$,

$$g_{i+1}(j^{-1}) \geq Rg_i(\rho). \tag{7.92}$$

Let $K = K(\rho)$ be the smallest integer in the range $1 \leq K \leq M + L$ such that $g_i(\rho) \leq f_i/2$ for all $K < i \leq M + L$, where we set $K = M + L$ if $g_{M+L}(\rho) > f_{M+L}/2$. We claim that there exists $\rho_1 > 0$ such that for all $\rho \leq \rho_1$,

$$f_K \leq 2g_K(\rho). \tag{7.93}$$

This follows by definition if $K \geq 2$. If $K = 1$, we set $\delta_0 := c_{10}/2$ and $\rho_1 := \rho_0(\delta_0, j)$ with c_{10} and ρ_0 from the first and second bullet points above respectively. By (7.91), for all $\rho \leq \rho_1$, we have $f_1 - g_1(\rho) \leq \delta_0$. This combined with (7.90) gives $f_1 - g_1(\rho) \leq f_1/2$ for all $\rho \leq \rho_1$, which implies (7.93).

In the following, for any $\delta > 0$, we let $\rho \leq \rho_0(\delta, j) \wedge \rho_1$, then both (7.91) and (7.93) hold. Furthermore, by (7.91), $f_i \leq 2\delta f_{i-1}$ for all $K < i \leq M + L$. Iterating this, we get

$$f_{M+L} \leq (2\delta)^{M+L-K} f_K \leq (2\delta)^{M+L-K} 2g_K(\rho), \tag{7.94}$$

where in the last inequality we have used (7.93). Repeated application of (7.92) gives that

$$g_K(\rho) \leq R(\rho)^{-1} R(j^{-1})^{K+1-M-L} g_{M+L}(j^{-1}). \quad (7.95)$$

Combining (7.94) with (7.95), we have

$$f_{M+L} \leq 2R(\rho)^{-1} R(j^{-1}) (2\delta/R(j^{-1}))^{M+L-K} g_{M+L}(j^{-1}).$$

Letting δ be sufficiently small such that $2\delta/R(j^{-1}) \leq 1$, we obtain (7.89) by setting $c(j)^{-1} = 2R(\rho)^{-1} R(j^{-1})$. This completes the proof. \square

References

- [1] Vincent Beffara and Pierre Nolin. On monochromatic arm exponents for 2D critical percolation. *Ann. Probab.*, 39(4):1286–1304, 2011.
- [2] I. Binder, L. Chayes, and H. K. Lei. On the rate of convergence for critical crossing probabilities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(2):672–715, 2015.
- [3] I. Binder and L. Richards. Convergence rates of random discrete model curves approaching sle curves in the scaling limit. *Preprint*, 2020.
- [4] Ilia Binder. Rate of convergence of critical interfaces to SLE curves. In *Extended Abstracts Fall 2019*, pages 43–50. Birkhäuser, 2021.
- [5] Federico Camia and Charles M Newman. Two-dimensional critical percolation: the full scaling limit. *Comm. Math. Phys.*, 268(1):1–38, 2006.
- [6] Hugo Duminil-Copin. Sixty years of percolation. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 2829–2856. World Scientific, 2018.
- [7] Hugo Duminil-Copin, Karol Kajetan Kozłowski, Dmitry Krachun, Ioan Manolescu, and Mendes Oulamara. Rotational invariance in critical planar lattice models. *ArXiv preprint arXiv:2012.11672*, 2020.
- [8] Christophe Garban, Gábor Pete, and Oded Schramm. Pivotal, cluster, and interface measures for critical planar percolation. *J. Amer. Math. Soc.*, 26(4):939–1024, 2013.

- [9] Geoffrey Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1999.
- [10] Nina Holden, Xinyi Li, and Xin Sun. Natural parametrization of percolation interface and pivotal points. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(1):7–25, 2022.
- [11] Richard Kenyon. The asymptotic determinant of the discrete laplacian. *Acta. Math.*, 185(2):239–286, 2000.
- [12] Harry Kesten. Scaling relations for 2D-percolation. *Comm. Math. Phys.*, 109(1):109–156, 1987.
- [13] Harry Kesten, Vladas Sidoravicius, and Yu Zhang. Almost all words are seen in critical site percolation on the triangular lattice. *Electron. J. Probab.*, 3(10):1–75, 1998.
- [14] Gregory Lawler, Xin Sun, and Wei Wu. Four-dimensional loop-erased random walk. *Ann. Probab.*, 47(6):3866–3910, 2019.
- [15] Gregory F. Lawler. The dimension of the frontier of planar Brownian motion. *Electron. Comm. Probab.*, 1(5):29–47, 1996.
- [16] Gregory F. Lawler, Oded Schramm, and Wendelin Werner. One-arm exponent for critical 2D percolation. *Electron. J. Probab.*, 7(3):1–13, 2002.
- [17] Xinyi Li and Daisuke Shiraishi. One-point function estimates for loop-erased random walk in three dimensions. *Electron. J. Probab.*, 24(111):1–46, 2019.
- [18] Dana Mendelson, Asaf Nachmias, and Samuel S. Watson. Rate of convergence for Cardy’s formula. *Comm. Math. Phys.*, 329(1):29–56, 2014.
- [19] Pierre Nolin. Near-critical percolation in two dimensions. *Electron. J. Probab.*, 13:1562–1623, 2008.
- [20] Larissa Richards. *Convergence Rates of Random Discrete Model Curves Approaching SLE Curves in the Scaling Limit*. PhD thesis, University of Toronto (Canada), 2021. Available for download at <https://hdl.handle.net/1807/106368>.
- [21] Oded Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, 118:221–288, 2000.

- [22] Oded Schramm. Conformally invariant scaling limits: an overview and a collection of problems. In *International Congress of Mathematicians. Vol. I*, pages 513–543. European Mathematical Society, 2007.
- [23] Oded Schramm and Jeffrey E. Steif. Quantitative noise sensitivity and exceptional times for percolation. *Ann. of Math.*, 171(2):619–672, 2010.
- [24] Stanislav Smirnov. Critical percolation in the plane: conformal invariance, Cardy’s formula, scaling limits. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(3):239–244, 2001.
- [25] Stanislav Smirnov and Wendelin Werner. Critical exponents for two-dimensional percolation. *Math. Res. Lett.*, 8(5-6):729–744, 2001.
- [26] Fredrik Johansson Viklund. Convergence rates for loop-erased random walk and other loewner curves. *Ann. Probab.*, 43(1):119–165, 2015.
- [27] Wendelin Werner. Lectures on two-dimensional critical percolation. In *IAS Park City Mathematics Series*, volume 16, pages 297–360. American Mathematical Society, 2009.

致谢

本科一路走来，过往所经历的一一点一滴塑造了今天的我，这一程值得感念的人和事太多太多。

首先感谢教导过，帮助过我的老师们，是他们的引领让我一步步走到了现在。感谢我的本科导师李欣意老师，他在我迷茫探询时引领我走向现代概率论，并让我逐渐坚定未来走学术道路的选择。他给予我细致的科研指导，让我从一个小白开始一点点上手做研究。他不仅是我的良师，亦是我的益友，在我人生中种种或大或小的选择上给予诚挚的建议，在我骄傲时加以鞭策，在我陷入低谷时施以鼓励。李老师一直以他豁达的人生观激励着我，让我不断发掘生活中新的可能性。感谢我的另一位科研导师丁剑老师，他亲力亲为地指导我的科研进阶之路，以他广博的知识储备以及富有创造力的洞见将我带到了新的学术高度。在他的引领与激励下，我逐渐培养了独立研究的能力，也发展出属于自己的研究视角。除开学术上的指导，他亦在许多关键问题上教会了我种种为人处事的深刻道理，给我带来了另一层意义下的成长。事实上，抛开令人叹服的数学能力，丁老师在与我一次次交流中所流露出的正直而善良的品格，方是他令我由衷敬佩的所在。我还要感谢每一位教授过我的老师，他们或为我打下了扎实的专业基础，或为我提供了学术探究的新思路。除开传道授业解惑之外，在这里我还想特别感谢概率方向的各位老师，他们每一位都热情而友善。感谢许惟钧老师在我对未来去向举棋不定时对我的激励，感谢章复熹老师在我科研起步时期与我做的讨论，感谢刘勇老师在随机分析课程上对我疑惑的解答以及在后续课程中给我的照顾。感谢丁老师，李老师和章老师在申请时给我做的鼎力推荐，让我能去到理想的地方继续追求学术梦想。感谢概率系的年轻老师们组织的种种活动，讨论班上的种种前沿话题极大地满足了我们的求知欲，讨论班后的茶歇与盒饭也满足了我们的胃口。正是老师们牵头的种种活动，构建了概率方向开放包容，和谐友善的交流环境，给了我家的感觉。还要感谢我认真负责的班主任李洵老师，她为我们创造了一个良好的班级环境，也为我争取到了许多荣誉。

感谢我的每一位合作者，从与他们的交流讨论中我学到了许多，无论是学术上还是学术之外。感谢李欣意老师将我引入渗流这一概率与物理交融的精彩领域，并带领我加入了第一个科研项目。感谢高一帆学长带我初上手科研，并包容我稚嫩的学术写作。感谢庄子杰学长为我树立的良好榜样，让我看到本科生科研的无限可能性。感谢丁剑老师在我尝试独立研究的初期费尽心思为我选题并孜孜不倦地与我讨论，无论是在学术视野，技术手段还是论文写作上都给了我极大的帮助。感谢巩舒阳学长毫无保留地投入到我们的合作中，让我在科研上拥有了并肩奋斗的战友。感谢常寅山老师与我在成都的讨论与合作，每一次与他交流都能带来新的想法。还要感谢我现在的合作者们，蔡振豪，刘昱，李章颂，黄润东，Brice Huang 以及 Mark Sellke，他们每一个人都为我提供了新的视角与想法，我相信在不远的将来我们会有共同的新工作产出。

感谢我的学长学姐们，他们在我探索的道路上无私地给予我帮助，让我看清前路的方向。感谢

卢维潇学长从我初中时期起就对我关照有加。他在大一大二时解答过我无数的专业相关问题，帮我打下良好的数学基础；他就像引路人，无数次为我提供或大或小的人生选择上的建议，我也将去往波士顿继续追随他的脚步。感谢成都七中在数院的其他学长学姐们，郭子棋，田翊，陈鸿宇，石元峰，姚舜天，彭湜，李为远，席国栋以及罗月桐，他们作为过来人给我的建议总是令我受益匪浅，而七中人共同的情结使我们即使天各一方也会紧密相连。感谢舒亦展学长在我选择方向时期给我的指引，感谢高一帆与庄子杰两位学者在我初入概率时对我的引导，感谢杨泓暾学长和冯煜阳学姐对我在包括但远不限于申请的多个方面的鼓励与建议，他们都给了我极大的帮助。特别感谢概率系的各位学长学姐们，蔡振豪，夏傲腾，刘昱，庄子杰，何天成，巩舒阳，孔繁湜，马恒，姜懿洋，刘泽霖，刘润声，蔡格非，杨鹏，梁渝涛，冯煜阳和张江昊，他们营造了概率系和谐而积极的学术与生活氛围，也带我快速融入了概率方向大家庭。他们自发组织的讨论班极大地拓宽了我的视野，也培养了我组织经营学生讨论班的责任心。

感谢我的同学与朋友们，他们给了我本科四年最多的陪伴。感谢我从初高中以来的挚友们，敖睿成，景虹皓，张遂初，徐苇杭，黄轶之，吴熙楠，郭维豪和陈博洋，与他们相处往往是我放松的时光。感谢我在数院遇到的各位志同道合的朋友，赵文浩，李章颂，范哲睿，刘浩宇，与他们的交流讨论令我受益匪浅，他们也为我的学术生活增色许多。感谢我的室友们，朱辰宇，马允轩和李咏璋，他们让我度过了四年和谐而欢乐的寝室时光。感谢我所遇到的每一个在我生命中留下痕迹的人，也许我们的交集很浅，抑或我们只能彼此陪伴一小段旅途，但我真诚地感念每一次的遇见，并祝福他们都拥有美好的前程。

我还要感谢我的父母，谢谢他们对我一如既往的支持与鼓励，也谢谢他们对我溢于言表的爱。经过大学四年的洗礼，我渐渐褪去青涩与稚嫩，一步步走向成熟，希望能早日报答他们的养育之恩。

最后我要感谢北京大学，感谢这里的环境与氛围将我重塑，让我成为了今天的自己。未名湖畔的风，博雅塔前的云，燕园里的赤橙黄绿，红楼飞雪，抚平了多少年少时期无处安放的思绪。园子里的种种人与事，让我一点点将“思想自由，兼容并包”这八个大字刻入骨髓，学会平等地看待并尊重多元的价值观，也发展出自由选择个性的勇气。这里严谨而轻松的学术氛围让我确认了未来的理想，也给了我付诸实际的力量。我始终记得入学典礼上物理学院高原宁院士对我们新生讲的话：“也许未来某一天你们会放弃理想而向现实妥协，但我希望那一天不要发生在北大”。如今即将踏出北大的校门，我可以骄傲地说，我依旧秉持着我的理想主义，而且比过去生命中的任何一个时期都更为坚定。感谢北大守护了我的理想，我不能确定以后的路会是怎样，但相信无论多少年后回首，我都会怀念这段纯真的时光，并将眼底那片未名水，胸中那一轮黄河月永远珍藏。

北京大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名：杜航 日期：2023年5月30日

学位论文使用授权说明

(必须装订在提交学校图书馆的印刷本)

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校 ☐ 一年 / ☐ 两年 / ☐ 三年以后，在校园网上全文发布。

(保密论文在解密后遵守此规定)

论文作者签名：杜航 导师签名：李锐
日期：2023年5月30日

