

THRESHOLDS FOR DETECTING AN ANOMALOUS PATH FROM NOISY ENVIRONMENTS

BY SHIRSHENDU CHATTERJEE^{*,1} AND OFER ZEITOUNI^{†,‡,2}

City University of New York, City College^{*}, Weizmann Institute of Science[†] and
New York University[‡]

We consider the “searching for a trail in a maze” composite hypothesis testing problem, in which one attempts to detect an anomalous directed path in a lattice 2D box of side n based on observations on the nodes of the box. Under the signal hypothesis, one observes independent Gaussian variables of unit variance at all nodes, with zero mean off the anomalous path and mean μ_n on it. Under the null hypothesis, one observes i.i.d. standard Gaussians on all nodes. Arias-Castro et al. [*Ann. Statist.* **36** (2008) 1726–1757] showed that if the unknown directed path under the signal hypothesis has known initial location, then detection is possible (in the minimax sense) if $\mu_n \gg 1/\sqrt{\log n}$, while it is not possible if $\mu_n \ll 1/\log n \sqrt{\log \log n}$. In this paper, we show that this result continues to hold even when the initial location of the unknown path is not known. As is the case with Arias-Castro et al. [*Ann. Statist.* **36** (2008) 1726–1757], the upper bound here also applies when the path is undirected. The improvement is achieved by replacing the linear detection statistic used in Arias-Castro et al. [*Ann. Statist.* **36** (2008) 1726–1757] with a polynomial statistic, which is obtained by employing a multiscale analysis on a quadratic statistic to bootstrap its performance. Our analysis is motivated by ideas developed in the context of the analysis of random polymers in Lacoin [*Comm. Math. Phys.* **294** (2010) 471–503].

1. Introduction.

1.1. *General problem.* In this paper, we will address the problem of detecting anomalous paths within a finite two-dimensional lattice, with an unknown start-point. We begin with describing the context for our results. Our presentation and motivation are strongly influenced by [1], to which we refer for additional background.

Suppose we are given a graph G with node set V and a random variable X_v attached to each node $v \in V$. We observe a realization of this process and wish to know whether all the variables at the nodes have the same behavior in the sense

Received May 2017; revised September 2017.

¹Supported in part by Simons Foundation Collaborative Research Funds.

²Supported in part by an Israel Science Foundation grant.

MSC2010 subject classifications. Primary 62C20, 62G10; secondary 82B20, 60K37.

Key words and phrases. Detecting a chain of nodes in a network, minimax detection, random scenery.

that they are all sampled independently from a common distribution F_0 (the null hypothesis, which in this paper will always be the standard Gaussian distribution), or whether there is a path in the network, that is, a chain of consecutive nodes connected by edges, along which the variables at the nodes, still independent of each other and of the variables off the path, have a different distribution F_1 (the signal hypothesis, which in this paper will always be the Gaussian distribution with nonzero mean and unit variance). This is thus a composite-hypothesis testing problem.

In this paper, as in [1], we focus on the case where G is a box V_n of side n in the two-dimensional Euclidean lattice, and the path under the signal hypothesis is a *directed* path. What distinguishes our analysis from the case treated in [1] is that we allow for an unknown starting point. Our main result (see Theorem 1.1 below) is that, similar to the case treated in [1], where the starting point is unknown, if the mean μ_n along the unknown path satisfies $\mu_n \geq C/\sqrt{\log n}$ for some large constant C , then detection is possible in the sense that a sequence of tests which are asymptotically powerful exists. (It follows from the main result in [1] that if $\mu_n \ll 1/\log n \sqrt{\log \log n}$, all tests are asymptotically powerless.) It is not hard to verify that our results concerning asymptotically powerful tests apply verbatim to the case of undirected *box crossing* paths, whose starting and ending points lie on two opposite sides of V_n , and to the case of undirected *annulus crossing* paths, whose starting point lies within a macroscopic sub-box of V_n .

The main difference between the analysis here and in [1] is in the test which is used for the hypothesis testing. In [1], one uses a test that is based on a linear statistics of the observations, where the weights are proportional to the inverse of the distance from the (known) initial point of the path. These tests clearly cannot be used in the case where the initial point is not known. Instead, in this paper we use tests that are based on quadratic and higher order polynomials of the observations, with nonhomogeneous weights. These are motivated by the success that certain quadratic forms had in the evaluation of the free energy of directed polymers in dimension $1 + 1$; see [5]. We note that a naive application of these quadratic test statistics leads to a detection threshold of order $1/(\log n)^{1/4}$ (see Section 2.1.2). The test we eventually use is based on a bootstrapped version of the simple quadratic form of the observations, whose analysis requires us to perform a multiscale analysis of somewhat modified detection problems.³

1.2. *Mathematical formulation of the detection problem.* In this section, we will formalize the detection problem. We consider the two-dimensional lattice $\mathbb{L}^2 = (\mathcal{V}^2, \mathcal{E}^2)$ with

$$(1.1) \quad \begin{aligned} \text{node set } \mathcal{V}^2 &:= \{\mathbf{x} \in \mathbb{Z}^2 : x_1 - x_2 \text{ is even}\} \quad \text{and} \\ \text{edge set } \mathcal{E}^2 &:= \{\langle \mathbf{x}, \mathbf{y} \rangle : \mathbf{x}, \mathbf{y} \in \mathcal{V}^2 \text{ and } \mathbf{x} \sim \mathbf{y}\}, \end{aligned}$$

³As pointed out by H. Lacoin, a similar in spirit analysis was used earlier in the random polymers context; see [2].

where $\mathbf{x} \sim \mathbf{y}$ if $x_1 \neq y_1, x_2 \neq y_2$ and $\|\mathbf{x} - \mathbf{y}\|_1 = 2$. We will use $\mathcal{H}_i := \{\mathbf{x} \in \mathbb{Z}^2 : x_1 = i\}$ to denote the i th hyperplane. We will consider a set of finite two-dimensional graphs \mathcal{G}_n , which consists of certain subgraphs of \mathbb{L}^2 induced by nodes in the hyperplanes $\bigcup_{0 \leq i < n} \mathcal{H}_i$. For $a \geq 0$, let $\mathcal{V}_n^{(a)}$ denote the node set

$$\mathcal{V}_n^{(a)} := \bigcup_{i=0}^{n-1} (\mathcal{H}_i \times [-i - an, i + an]) \quad \text{and} \quad \mathcal{E}_n^{(a)} := (\mathcal{V}_n^{(a)}, \mathcal{E}^2|_{\mathcal{V}_n^{(a)}})$$

be the subgraph of \mathbb{L}^2 induced by nodes in $\mathcal{V}_n^{(a)}$. We also introduce the notation

$$\begin{aligned} \mathcal{Z}_n^{(a)} &:= \{\mathbf{z} \in \mathbb{Z}^2 : 0 \leq z_1 < n, |z_2| \leq an + z_1\}, \\ [n] &:= \{1, 2, \dots, n\}, [n]_0 := \{0, 1, \dots, n - 1\} \end{aligned}$$

and

$$\mathcal{G}_n := \{\mathcal{G}_n^{(a)} : a \geq 0\}.$$

Having defined the family of two-dimensional finite graphs, we define the collection of semi- directed nearest-neighbor paths (left to right crossing) on these graphs, with a starting point on the hyperplane \mathcal{H}_0 and endpoint on \mathcal{H}_{n-1} . For a graph $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{G}_n$, let

$$(1.2) \quad \begin{aligned} \mathcal{P}(\mathcal{G}_n) &:= \{\boldsymbol{\pi} = \langle \boldsymbol{\pi}_0, \dots, \boldsymbol{\pi}_{n-1} \rangle : \boldsymbol{\pi}_i \in \mathcal{H}_i \\ &\text{for all } i \in [n]_0 \text{ and } \boldsymbol{\pi}_i \sim \boldsymbol{\pi}_{i-1} \text{ for all } i \in [n - 1]\}; \end{aligned}$$

see Figures 1 and 2 for an instance of such a path on graphs in $\mathcal{G}_n^{(a)}$, $a = 0$ and $a > 0$. In particular, the starting point of the collection of paths $\mathcal{P}(\mathcal{G}_n)$ is known if $\mathcal{G}_n = \mathcal{G}_n^{(0)}$ and unknown if $\mathcal{G}_n = \mathcal{G}_n^{(a)}$ with $a > 0$.

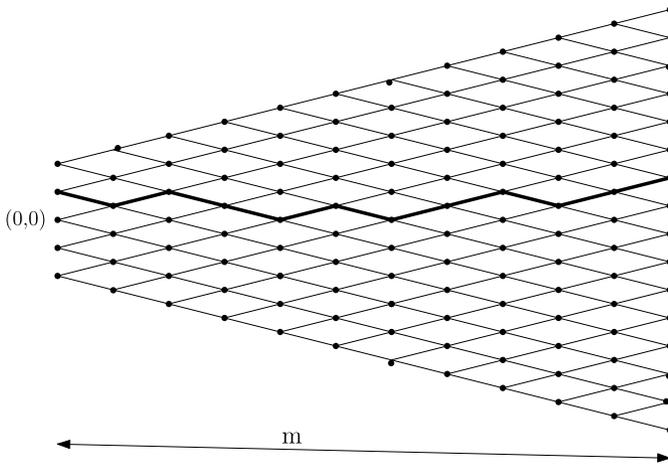


FIG. 1. This is a picture of a graph in $\mathcal{G}_n^{(a)}$, $a > 0$. The bold line represents a path in the two-dimensional finite lattice with unknown initial location.

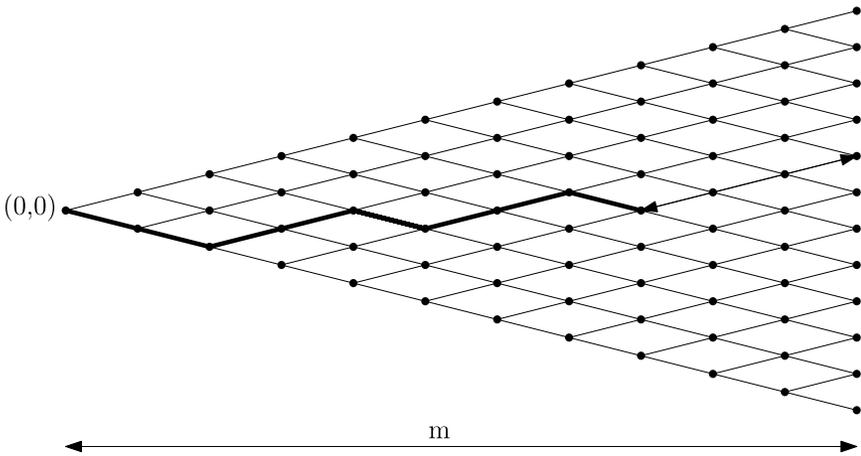


FIG. 2. This is a picture of a graph in $\mathcal{G}_n^{(0)}$. The bold line represents a path in the two-dimensional finite lattice with known initial location.

We next introduce the statistical hypothesis problem on $(\mathcal{G}_n, \mathcal{P}(\mathcal{G}_n))$. To each node \mathbf{v} of the graph \mathcal{G}_n , one attaches a random variable $X_{\mathbf{v}}$. We assume that all random variables are independent and consider the following hypothesis testing problem:

- *Null hypothesis H_0* : The random variables $\{X_{\mathbf{v}} : \mathbf{v} \in \mathcal{V}_n\}$ are i.i.d. with common distribution $N(0, 1)$.
- *Alternate (signal) hypothesis $H_{1,\pi}$* : it is a composite hypothesis $\bigcup_{\pi \in \mathcal{P}(\mathcal{G}_n)} H_{1,\pi}$, where, under $H_{1,\pi}$, the random variables $\{X_{\mathbf{v}} : \mathbf{v} \in \mathcal{V}_n\}$ are independent with

$$X_{\mathbf{v}} \stackrel{d}{=} \begin{cases} N(\mu_n, 1) & \text{if } \mathbf{v} \in \pi, \\ N(0, 1) & \text{otherwise,} \end{cases} \quad \text{for some } \mu_n > 0.$$

In other words, the null hypothesis is that the random variables $\{X_{\mathbf{v}}\}$ represent a random scenery, whereas the alternative hypothesis suggests that there is an anomalous path along which the mean of the random variables are nontrivial. We refer to the above hypothesis testing problem as the “ $(\mathcal{P}(\mathcal{G}_n), \mu_n, \Phi)$ detection problem on \mathcal{G}_n .” Φ represents the cdf of $N(0, 1)$.

The detection threshold is the minimum value of $\mu = \mu_n$ for which one can reliably decide whether or not there is an anomalous path which does not follow the null distribution. The threshold depends on the criterion used for judging the performance of the decision rule. There are mainly two paradigms in statistical decision theory, namely the Bayesian and the minimax approach. We will consider the second approach. Recall that a nonrandomized test T_n is a measurable function of the collection of random variables $(X_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$ taking values in $\{0, 1\}$. The minimax risk of such a test T_n is defined as

$$\gamma(T_n) := \mathbb{P}_0(\text{Type I error}) + \sup_{\pi \in \mathcal{P}(\mathcal{G}_n)} \mathbb{P}_{1,\pi}(\text{Type II error}),$$

where

$$\mathbb{P}_0(\text{Type I error}) = \mathbb{P}_0(T_n = 1) \quad \text{and} \quad \mathbb{P}_{1,\pi}(\text{Type II error}) = \mathbb{P}_{1,\pi}(T_n = 0).$$

Here and later, \mathbb{P}_0 denotes the probability distribution under the null hypothesis and $\mathbb{P}_{1,\pi}$ denotes the probability distribution under the alternative hypothesis when $\pi \in \mathcal{P}(\mathcal{G}_n)$ is the anomalous path. A sequence of tests $\{T_n\}_{n \geq 1}$ for the hypothesis testing problem $(\mathcal{P}(\mathcal{G}_n), \mu_n)$ will be called asymptotically powerful if

$$\lim_{n \rightarrow \infty} \gamma(T_n) = 0,$$

and it will be called asymptotically powerless if

$$\lim_{n \rightarrow \infty} \gamma(T_n) \geq 1.$$

1.3. *Main result.* The main result of this paper is the following theorem.

THEOREM 1.1. *Fix $a \geq 0$. There is a finite constant C large enough such that for any sequence of means $\{\mu_n\}_{n \geq 1}$ satisfying $\mu_n \sqrt{\log n} \geq C$, there exists a sequence of tests $\{T_n\}_{n \geq 1}$ for the hypothesis testing problem $(\mathcal{P}(\mathcal{G}_n^{(a)}), \mu_n, \Phi)$ which is asymptotically powerful. On the other hand, for any sequence of means $\{\mu_n\}_{n \geq 1}$ satisfying $\mu_n \log n \sqrt{\log \log n} \rightarrow 0$ as $n \rightarrow \infty$, all sequence of tests $\{T_n\}_{n \geq 1}$ for the hypothesis testing problem $(\mathcal{P}(\mathcal{G}_n^{(a)}), \mu_n, \Phi)$ will be asymptotically powerless.*

REMARK 1.2. The case $a = 0$ of Theorem 1.1 is contained in [1].

REMARK 1.3. The asymptotically powerful part of the assertion of Theorem 1.1 holds for the detection problem $(\tilde{\mathcal{P}}(\mathcal{G}_n^{(a)}), \mu_n, \Phi)$, where $\tilde{\mathcal{P}}(\mathcal{G}_n^{(a)})$ consists of undirected paths on $\mathcal{G}_n^{(a)}$ having their one endpoint in \mathcal{H}_0 and the other endpoint in \mathcal{H}_{n-1} .

REMARK 1.4. For $0 < b < a$, the assertion of Theorem 1.1 holds for the detection problem $(\mathcal{P}(a, b), \mu_n, \Phi)$, where $\mathcal{P}(a, b)$ consists of directed paths in the subgraph of \mathbb{L}^2 induced by $[-an, an]^2 \cap \mathcal{V}^2$ having their one endpoint in $[-bn, bn]^2 \cap \mathcal{V}^2$ and the other endpoint on the boundary of $[-an, an]^2 \cap \mathcal{V}^2$.

1.4. *Notation.* Throughout the paper, we will use the following notation:

- For $n \in \mathbb{N}$, we will use $[n]$ to denote the set $\{1, 2, \dots, n\}$.
- For $\mathbf{M} \in \mathbb{R}^{m \times n}$, we will use $\|\mathbf{M}\|_F$ to denote the Frobenius norm

$$\|\mathbf{M}\|_F := \sqrt{\text{Trace}(\mathbf{M}^T \mathbf{M})} = \sqrt{\sum_{i \in [m], j \in [n]} M_{i,j}^2}.$$

Finally, for sequences $a = (a_n)$ and $b = (b_n)$ we write $a \asymp b$ to mean that there exists a finite universal constant $c > 0$ so that $1/c \leq a_n/b_n \leq c$ for all n large.

2. Proof of Theorem 1.1.

2.1. *Upper bound for the detection threshold.* In this section, we will eventually show that a sequence of asymptotically powerful tests exists for the hypothesis testing problem H_0 versus H_1 if $\mu_n \sqrt{\log n} > C$ for some large enough constant C . We present the proof of a weaker version of this assertion in Proposition 2.2, which is then bootstrapped in conjunction with a renormalization argument to complete the proof (see Section 2.1.5). First, we need to introduce certain quadratic forms, which play a crucial role in the proof.

2.1.1. *Quadratic forms associated with the detection problem.* We next introduce some useful notation. In order to arrange the vertices of \mathcal{V}_n and associated random variables in an order, we define the following partial order:

For $\mathbf{x}, \mathbf{y} \in \mathcal{V}^2$, define $\mathbf{x} \preceq \mathbf{y}$ if either $x_1 < y_1$ or $x_1 = y_1$ and $x_2 < y_2$.

Using this partial order, we order the random variables $\{X_{\mathbf{v}} : \mathbf{v} \in \mathcal{V}_n\}$ accordingly to have the $|\mathcal{V}_n| \times 1$ column vector \mathbf{X}_n . For $A \subset \mathcal{V}_n$, we use $\mathbb{1}_A$ to denote the $|\mathcal{V}_n| \times 1$ column vectors defined by

$$\mathbb{1}_A(\mathbf{v}) = \begin{cases} 1 & \text{if } \mathbf{v} \in A, \\ 0 & \text{otherwise,} \end{cases} \quad \text{for } \mathbf{v} \in \mathcal{V}_n.$$

In order to describe the test T_n that will separate H_0 and H_1 , we also need the following equivalence relation:

For $\mathbf{x}, \mathbf{y} \in \mathcal{V}^2$, define $\mathbf{x} \rightsquigarrow \mathbf{y}$ if $x_1 \neq y_1$ and $|x_2 - y_2| \leq |x_1 - y_1|$.

It is easy to see that the above is an equivalence relation. For this equivalence relation and partial order described above, we write

$$[\mathbf{x}] := \{\mathbf{y} \in \mathcal{V}_n : \mathbf{x} \rightsquigarrow \mathbf{y}\} \quad \text{and} \quad \mathbf{x} \preceq \mathbf{y} \quad \text{if } \mathbf{x} \rightsquigarrow \mathbf{y} \text{ and } \mathbf{x} \preceq \mathbf{y}.$$

Using the above partial order and equivalence relation, we define the $|\mathcal{V}_n| \times |\mathcal{V}_n|$ matrix $[\mathbf{A}(\mathcal{V}_n)]$ associated with the full vertex set \mathcal{V}_n by

$$(2.1) \quad [\mathbf{A}(\mathcal{V}_n)] = ([A(\mathcal{V}_n)]_{\mathbf{x},\mathbf{y}})_{\mathbf{x},\mathbf{y} \in \mathcal{V}_n} \quad \text{where } [A(\mathcal{V}_n)]_{\mathbf{x},\mathbf{y}} = \frac{1}{|x_1 - y_1|} \mathbf{1}_{\{\mathbf{x} \rightsquigarrow \mathbf{y}\}},$$

$$[\bar{\mathbf{A}}(\mathcal{V}_n)] := (\sqrt{2} \|\mathbf{A}(\mathcal{V}_n)\|_F)^{-1} [\mathbf{A}(\mathcal{V}_n)].$$

The matrix $\mathbf{A}(\mathcal{V}_n)$ will play a special role in our argument. The following lemma, whose proof is postponed to Section 3.1, collects some of its elementary properties. In the lemma and its proof, we write $\mathcal{P}_n = \mathcal{P}(\mathcal{G}_n)$.

LEMMA 2.1. *For any graph $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{G}_n$ and for the matrix $[\mathbf{A}(\cdot)]$ as defined in (2.1):*

- (1) $\|[\mathbf{A}(\mathcal{Y}_n)]\|_F \asymp n\sqrt{\log n}$,
- (2) $\|[\mathbf{A}(\mathcal{Y}_n)]\| = O(n)$, so $\|[\bar{\mathbf{A}}(\mathcal{Y}_n)]\| = O(1/\sqrt{\log n})$,
- (3) $\mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \mathbb{1}_\pi \asymp n \log n$ for any $\pi \in \mathcal{P}_n$,
- (4) $\mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)]^2 \mathbb{1}_\pi \asymp n^2$ for any $\pi \in \mathcal{P}_n$,
- (5) $\mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \text{Diag}(\mathbb{1}_\pi) [\mathbf{A}(\mathcal{Y}_n)] \mathbb{1}_\pi \asymp n(\log n)^2$ for any $\pi \in \mathcal{P}_n$,
- (6) $\|[\mathbf{A}(\mathcal{Y}_n)]^2\|_F \asymp n^4 \log n$, so $\|[\bar{\mathbf{A}}(\mathcal{Y}_n)]^2\|_F \asymp (\log n)^{-1/2}$.

Note that Lemma 2.1 describes properties of the matrix $[\mathbf{A}(\mathcal{Y}_n)]$.

2.1.2. A weaker version of Theorem 1.1.

PROPOSITION 2.2. *In the set up of Theorem 1.1, if $\mu_n(\log n)^{1/4} \rightarrow \infty$ as $n \rightarrow \infty$, there is a sequence of asymptotically powerful tests for the hypothesis testing problem H_0 versus H_1 .*

PROOF. Let \mathbf{Z}_n be a $|\mathcal{Y}_n| \times 1$ column vector consisting of i.i.d. $N(0, 1)$ random variables. Consider the quadratic form $\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n$, where $[\mathbf{A}(\mathcal{Y}_n)]$ is the matrix defined in (2.1). Since $[\mathbf{A}(\mathcal{Y}_n)]$ has zero diagonal entries, each summand of $\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n$ and $Z_i (\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n)$, $i \in [|\mathcal{Y}_n|]$, has mean 0, as all of them are product of independent random variables having mean 0. So

$$(2.2) \quad \mathbb{E} \mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n = 0 \quad \text{and} \quad \mathbb{E} [(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n) \mathbf{Z}_n] = \mathbf{0}.$$

Also, noting that the summands $\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n$ are uncorrelated, and using (1) of Lemma 2.1,

$$(2.3) \quad \begin{aligned} \mathbb{E} [(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n)^2] &= 4 \sum_{(i,u) \in \mathcal{Y}_n} \sum_{(j,v) \in \mathcal{Y}_n: (i,u) \prec (j,v)} (j-i)^{-2} \\ &= 2 \|[\mathbf{A}(\mathcal{Y}_n)]\|_F^2 \asymp n^2 \log n. \end{aligned}$$

Now, using the partial order \prec , we order the random variables $\{X_v : v \in \mathcal{Y}_n\}$ attached to the nodes to have the $|\mathcal{Y}_n| \times 1$ column vector \mathbf{X}_n . Define the quadratic form $Q_n := \mathbf{X}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{X}_n$ and the test $T_n := \mathbf{1}_{\{Q_n > \mu_n^2 n \log n / 2\}}$. In order to compute $\gamma(T_n)$ note that

$$\mathbf{X}_n \stackrel{d}{=} \begin{cases} \mathbf{Z}_n & \text{under } H_0, \\ \mathbf{Z}_n + \mu_n \mathbb{1}_\pi & \text{under } H_{1,\pi}, \end{cases}$$

so using (2.2), (2.3) and Lemma 2.1 we get

$$\begin{aligned} \mathbb{E}_0 Q_n &= \mathbb{E} \mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n = 0, \\ \mathbb{E}_{1,\pi} Q_n &= \mathbb{E} \mathbf{Z}_n^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n + 2\mu_n \mathbb{E} \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \mathbf{Z}_n + \mu_n^2 \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \mathbb{1}_\pi \\ &= \mu_n^2 \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \mathbb{1}_\pi \asymp \mu_n^2 n \log n, \end{aligned}$$

$$\begin{aligned} \text{Var}_0 Q_n &= \mathbb{E}[(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n)^2] \asymp n^2 \log n, \\ \text{Var}_{1,\pi} Q_n &= \text{Var}(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n + 2\mu_n \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n) = \text{Var}(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n) \\ &\quad + \text{Var}(2\mu_n \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n) \\ &\quad + 2 \text{Cov}(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n, 2\mu_n \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n) \\ &= \mathbb{E}[(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n)^2] + 4\mu_n^2 \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \mathbb{E}(\mathbf{Z}_n \mathbf{Z}_n^T) [\mathbf{A}(\mathcal{V}_n)] \mathbb{1}_\pi \\ &\quad + 4\mu_n \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \mathbb{E}(\mathbf{Z}_n \mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n) \\ &= \mathbb{E}[(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n)^2] + 4\mu_n^2 \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)]^2 \mathbb{1}_\pi \\ &\asymp n^2 \log n + \mu_n^2 n^2. \end{aligned}$$

Using the above estimates and Chebychev’s inequality,

$$\begin{aligned} \mathbb{P}_0(T_n = 1) &\leq \mathbb{P}_0\left(|Q_n - \mathbb{E}_0 Q_n| > \frac{1}{2} \mu_n^2 n \log n\right) \\ &\leq 4 \frac{\text{Var}_0(Q_n)}{(\mu_n^2 n \log n)^2} \leq c \frac{n^2 \log n}{(\mu_n^2 n \log n)^2} = \frac{c}{\mu_n^4 \log n} \quad \text{and} \\ \mathbb{P}_{1,\pi}(T_n = 0) &\leq \mathbb{P}_{1,\pi}\left(|Q_n - \mathbb{E}_{1,\pi} Q_n| \geq \frac{1}{2} \mu_n^2 n \log n\right) \\ &\leq 4 \frac{\text{Var}_{1,\pi}(Q_n)}{(\mu_n^2 n \log n)^2} \leq c \frac{n^2 \log n + \mu_n^2 n^2}{(\mu_n^2 n \log n)^2} \\ &= \frac{c}{\mu_n^4 \log n} + \frac{c}{\mu_n^2 \log^2 n} \end{aligned}$$

for some constant c . Since the upper bounds in the above display are $o(1)$, we see that $\lim_{n \rightarrow \infty} \gamma(T_n) = 0$. This completes the proof. \square

Proposition 2.2 gives a weak upper bound for the detectability threshold. In order to improve this bound, we will use a renormalization argument. In order to employ our renormalization argument, we need to generalize the detection problem described in the **Introduction**. We define the necessary generalization step by step in the following section.

2.1.3. *Generalized detection problem.* Recall from Section 1.2 that $\mathcal{H}_i = i + 2\mathbb{Z}$ denotes the i th hyperplane of \mathbb{L}^2 . We extend the notion of a hyperplane by including unordered pairs of consecutive nodes from the same hyperplane. For a graph $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$, define the associated *generalized hyperplanes* as follows:

$$\begin{aligned} \widetilde{\mathcal{H}}_i(\mathcal{G}_n) &:= \{\{\mathbf{v}\} : \mathbf{v} \in \mathcal{V}_n \text{ and } v_1 = i\} \\ &\cup \{\{\mathbf{u}, \mathbf{v}\} : \mathbf{u}, \mathbf{v} \in \mathcal{V}_n, v_1 = u_1 = i \text{ and } |v_2 - u_2| = 2\}, \\ &i \in [n]_0. \end{aligned}$$

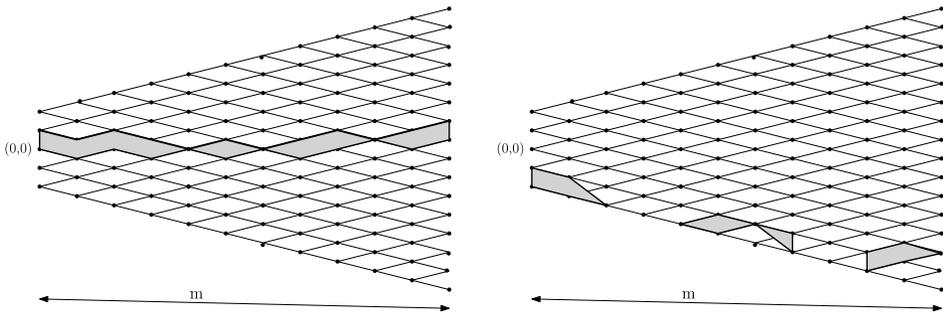


FIG. 3. This is a picture of a graph in $\mathcal{G}_n^{\text{unknown}}$. The shaded region represents a generalized path in the two-dimensional finite lattice with unknown initial location.

So each $\tilde{\mathcal{H}}_i(\mathcal{G}_n)$ consists of singletons and doubletons. We extend the definition of neighboring relationship \sim defined in (1.1) to a new relation \sim on $\bigcup_{i \in [n]_0} \tilde{\mathcal{H}}_i(\mathcal{G}_n)$. We say that $A \in \tilde{\mathcal{H}}_i(\mathcal{G}_n)$ and $B \in \tilde{\mathcal{H}}_j(\mathcal{G}_n)$ are neighbors, that is,

$$A \sim B \quad \text{if } a \sim b \text{ for some } a \in A \text{ and } b \in B.$$

In the same spirit, a *generalized path* Π on \mathcal{G}_n will be union of finite sequences of successive neighbors from $\bigcup_{i \in [n]_0} \tilde{\mathcal{H}}_i(\mathcal{G}_n)$. A generalized path may be incomplete in the sense that it may not intersect all hyperplanes. Define

$$\begin{aligned} \tilde{\mathcal{P}}(\mathcal{G}_n) := \{ \Pi = \langle \Pi_0, \dots, \Pi_{n-1} \rangle : \text{for all } i \in [n]_0 \text{ either } \Pi_i = \emptyset \text{ or } \Pi_i \in \tilde{\mathcal{H}}_i(\mathcal{G}_n), \\ \text{and } \Pi_i \sim \Pi_{i-1} \text{ whenever } \Pi_i, \Pi_{i-1} \neq \emptyset \}. \end{aligned}$$

See Figure 3 for a picture of such a generalized path for a graph in \mathcal{G}_n . In the generalized detection problem, we also assume that each node \mathbf{v} of the graph \mathcal{G}_n has a (observable) random variable $X_{\mathbf{v}}$ associated with it, and the random variables $\{X_{\mathbf{v}}\}$ are independent. We will refer to the collection of random variables $(X_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$ as *observables*. Suppose $(\nu_{\mathbf{v}, \Pi} \in \mathbb{R}_+, \mathbf{v} \in \mathcal{V}_n, \Pi \in \tilde{\mathcal{P}}(\mathcal{G}_n))$ is a collection of *signals* and $(Z_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$ and $(Y_{\mathbf{v}, \Pi}, \mathbf{v} \in \mathcal{V}_n, \Pi \in \tilde{\mathcal{P}}(\mathcal{G}_n))$, which we will refer to as *basic noise* and *additional noise*, respectively, are two collections of (possibly unobservable) random variables satisfying the following properties.

PROPERTY 2.3. The noise variables satisfy the following:

1. The random variables $((Z_{\mathbf{v}}, Y_{\mathbf{v}, \Pi}), \mathbf{v} \in \mathcal{V}_n)$ are independent.
2. $Y_{\mathbf{v}, \Pi}$ has mean zero and variance at most 1.
3. For each $\mathbf{v} \in \mathcal{V}_n$, the random variables $Z_{\mathbf{v}}$ and $Y_{\mathbf{v}, \Pi}$ are uncorrelated, although they can be dependent.

Based on the signals, basic noise and additional noise variables as described above, we consider the following two hypotheses:

- Null hypothesis H_0 : $X_{\mathbf{v}} = Z_{\mathbf{v}}$ for all $\mathbf{v} \in \mathcal{V}_n$.
- Alternate (signal) hypothesis $H_{1,\Pi}$: it is a composite hypothesis $\bigcup_{\Pi \in \tilde{\mathcal{P}}(\mathcal{G}_n)} H_{1,\Pi}$, where, under $H_{1,\Pi}$,

$$(2.4) \quad X_{\mathbf{v}} = \begin{cases} Z_{\mathbf{v}} + Y_{\mathbf{v},\Pi} + \nu_{\mathbf{v},\Pi} & \text{if } \mathbf{v} \in \bigcup_{i=0}^{n-1} \Pi_i, \\ Z_{\mathbf{v}} & \text{otherwise.} \end{cases}$$

We refer to this hypothesis testing problem as “generalized detection problem.”

The following proposition summarizes the relevant properties of a certain quadratic statistic under the hypotheses of a generalized detection problem. Proposition 2.4 will play an important role in proving the upper bound in Theorem 1.1.

PROPOSITION 2.4. Fix $a \geq 0$. Let $\mathcal{G}_n^{(a)} = (\mathcal{V}_n, \mathcal{E}_n)$ be any graph in \mathcal{G}_n , $\{F_{\mathbf{v}}\}_{\mathbf{v} \in \mathcal{V}_n}$ be any collection of distributions on \mathbb{R} having mean 0, variance 1 and finite third moment. Consider the generalized hypothesis testing problem with observables $(X_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$, where the basic noise variables $(Z_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$, the additional noise variables $(Y_{\mathbf{v},\Pi}, \mathbf{v} \in \mathcal{V}_n, \Pi \in \tilde{\mathcal{P}}(\mathcal{G}_n))$ and the signals $(\nu_{\mathbf{v},\Pi}, \mathbf{v} \in \mathcal{V}_n, \Pi \in \tilde{\mathcal{P}}(\mathcal{G}_n))$ satisfy Property 2.3 and $Z_{\mathbf{v}}$ has distribution $F_{\mathbf{v}}$. Let $Q_n := Q_n[[\bar{\mathbf{A}}(\mathcal{V}_n)], (X_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)]$ and $W_n := W_n[[\bar{\mathbf{A}}(\mathcal{V}_n)], (Z_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)]$ be the quadratic forms based on the observables and basic noise variables, respectively, where $[\bar{\mathbf{A}}(\cdot)]$ is as in (2.1):

1. Let $\beta_3 := \max_{\mathbf{v} \in \mathcal{V}_n} \mathbb{E}(|Z_{\mathbf{v}}|^3)$. Then there exists a constant $C = C(a)$ so that

$$(2.5) \quad \sup_{x \in \mathbb{R}} |\mathbb{P}(W_n \leq x) - \Phi(x)| \leq C \left[(\log n)^{-1/2} + \left(\frac{\beta_3}{n} + \frac{\beta_3^2}{n \log^{3/2} n} \right)^{1/4} \right].$$

2. Let $\bar{\nu} := \max_{\mathbf{v} \in \Pi} \nu_{\mathbf{v},\Pi}$. There is a random variable U_n , satisfying $\mathbb{E}U_n = \mathbb{E}U_n W_n = 0$ and $\mathbb{E}U_n^2 \leq C\bar{\nu}^2 / \log n$, such that

$$Q_n \stackrel{d}{=} \begin{cases} W_n & \text{under } H_0, \\ W_n + U_n + \nu(\mathcal{G}_n) & \text{under } H_{1,\Pi}, \end{cases}$$

where $\nu(\mathcal{G}_n) := \sum_{\mathbf{v}, \mathbf{v}' \in \mathcal{V}_n} \nu_{\mathbf{v},\Pi} \nu_{\mathbf{v}',\Pi} [\bar{\mathbf{A}}(\mathcal{V}_n)]_{\mathbf{v}, \mathbf{v}'}$.

3. For $I \subset [n]$, let $\underline{\nu}_I := \min_{i \in I} \max_{\mathbf{v} \in \Pi_i} \nu_{\mathbf{v},\Pi}$. There is a constant $c = c(a) \in (0, 1)$ such that

$$(2.6) \quad \begin{aligned} c \underline{\nu}_I^2 &\leq \frac{\nu(\mathcal{G}_n)}{\sqrt{\log n}} \\ &\leq c^{-1} \bar{\nu}^2 \quad \text{for any } I \subset \{i \in [n]_0 : \Pi_i \neq \emptyset\} \text{ satisfying } |I| \geq n/2. \end{aligned}$$

In the above setup, we will interpret W_n , U_n and $v(\mathcal{G}_n)$ as the basic noise variable, additional noise variable and signal, for the graph \mathcal{G}_n , respectively.

The proof of Proposition 2.4 uses the properties of the matrix $\mathbf{A}(\mathcal{V}_n)$ contained in Lemma 2.1. The proof is postponed to Section 3.3. The last ingredients that we need for proving Theorem 1.1 are some distributional properties of quadratic forms, which we present in the following section.

2.1.4. *Moment bounds and Gaussian approximation for quadratic forms.* Let $(X_j, j \in [n])$ denote independent random variables such that $\mathbb{E}X_j = 0$ and $\mathbb{E}X_j^2 = 1$ for all $j \geq 1$. Let $\mathbf{A} = \{a_{j,k}\}_{j,k=1}^n \in \mathbb{R}^{n \times n}$ be such that

$$(2.7) \quad \begin{aligned} \mathbf{A} \text{ is a symmetric matrix,} \quad & a_{j,j} = 0 \quad \text{for all } j \in [n] \quad \text{and} \\ \text{Trace}(\mathbf{A}^2) &= \frac{1}{2}. \end{aligned}$$

Consider the quadratic forms

$$(2.8) \quad Q_n[\mathbf{A}, (X_j, j \in [n])] := \sum_{j,k=1}^n a_{j,k} X_j X_k \quad \text{and} \quad G_n(\mathbf{A}) := \sum_{j,k=1}^n a_{j,k} Y_j Y_k,$$

where $(Y_j, j \in [n])$ are i.i.d. with common distribution $N(0, 1)$. Keeping in mind that we will need upper bounds for the third moment of certain quadratic forms, we state the following moment estimate.

THEOREM 2.5 (Theorem 2 of [8]). *If $\mathbf{A} \in \mathbb{R}^{n \times n}$ satisfies (2.7) and $(X_j, j \in [n])$ are independent random variables having zero mean, then*

$$\mathbb{E}(|Q_n[\mathbf{A}, (X_j, j \in [n])]|^s) \leq 2^{5s/2} \Gamma(s/2 + 1/2) (\Gamma(s + 1/2))^{1/2} \max_{j \in [n]} \mathbb{E}(X_j^{2s}).$$

Other than moments, we will also need error bounds for Gaussian approximation of quadratic forms. In this context, Rotar’ [6] proved that under sufficiently weak conditions on the matrix \mathbf{A} and for large n , the distribution of $Q_n[\mathbf{A}, (X_j, j \in [n])]$ is close to that of $G_n(\mathbf{A})$. Gamkrelidze and Rotar’ [3] obtained bounds for the error of this approximation, which were improved by Rotar’ and Shervashidze [7]. Here is their result:

$$(2.9) \quad \begin{aligned} \text{Let } F_j(x) &= \mathbb{P}(X_j \leq x), \quad \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy, \\ v_j &:= 3 \int_{-\infty}^{\infty} x^2 |F_j(x) - \Phi(x)| dx, \\ s_j^2 &:= \sum_{k=1}^n a_{n,j,k}^2, \quad L := \sum_{j=1}^n v_j s_j^3 + \sum_{j,k=1}^n v_j v_k |a_{n,j,k}|^3, \\ \Delta &:= \text{Trace}(\mathbf{A}^4). \end{aligned}$$

THEOREM 2.6 (See [7]). *Assume that (2.7) holds and L, Δ are as in (2.9). If $\Delta < 1/2$, then there is an absolute constant C such that*

$$\begin{aligned} \sup_{x \in \mathbb{R}} |\mathbb{P}(Q_n[\mathbf{A}, (X_j, j \in [n])] \leq x) - \mathbb{P}(G_n(\mathbf{A}) \leq x)| \\ \leq C(1 - \log(1 - 2\Delta))^{3/4} L^{1/4}. \end{aligned}$$

Later, Götze and Tikhomirov [4] obtained an improved bound for the Kolmogorov distance between normalized quadratic forms of i.i.d. Gaussian random variables and the Gaussian distribution.

THEOREM 2.7 (Theorem 1 of [4]). *Assume that (2.7) holds and let $G_n(\cdot)$ be as in (2.8). Then*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(G_n(\mathbf{A}) \leq x) - \Phi(x)| \leq C \|\mathbf{A}\| \quad \text{for some absolute constant } C.$$

We will apply Theorems 2.6 and 2.7 to bound the Kolmogorov distance between the quadratic form $Q_n[[\bar{\mathbf{A}}(\mathcal{V}_n)], (U_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)]$, where $[\bar{\mathbf{A}}(\cdot)]$ is defined in (2.1), and the Gaussian distribution.

PROPOSITION 2.8. *Fix $a \geq 0$. For any graph $\mathcal{G}_n := \mathcal{G}_n^{(a)} = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{G}_n$ and any collection of random variables $(U_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$ having zero mean, unit variance and finite third moment, if $[\bar{\mathbf{A}}(\mathcal{V}_n)]$ is the matrix as defined in (2.1), $Q_n = Q_n[[\bar{\mathbf{A}}(\mathcal{V}_n)], (U_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)]$ is the quadratic form as defined in (2.8) and $\beta_3 := \max_{\mathbf{v} \in \mathcal{V}_n} \mathbb{E}(|U_{\mathbf{v}}|^3)$, then*

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(Q_n \leq x) - \Phi(x)| \leq C \left[(\log n)^{-1/2} + \left(\frac{\beta_3}{n} + \frac{\beta_3^2}{n \log^{3/2} n} \right)^{1/4} \right],$$

where $C = C(a)$ is an absolute constant.

The proof of Proposition 2.8 uses properties of the matrix $\mathbf{A}(\mathcal{V}_n)$ contained in Lemma 2.1. The proof is postponed to Section 3.2.

2.1.5. Proof of the upper bound for the detectability threshold. Proposition 2.4 will play a crucial role in the proof of Theorem 1.1. It will be used in conjunction with a renormalization argument.

PROOF OF THEOREM 1.1, UPPER BOUND. Throughout the proof, we fix $a \geq 0$ and write $\mathcal{G}_n = \mathcal{G}_n^{(a)}$. Let $n_0 := n$, $n_k := n_{k-1}^{\varepsilon_k}$ for $k \in [K + 1]$, where $K \in \mathbb{N}$ and $\varepsilon_1, \dots, \varepsilon_K \in (0, 1)$ will be specified later [see (2.13) and (2.16)], and set $\varepsilon_{K+1} := 0$.

We construct a sequence of hierarchical partition $\{\mathcal{B}_k\}_{0 \leq k \leq K}$ of \mathcal{G}_n . \mathcal{B}_k will consist of vertex-disjoint subgraphs of \mathcal{G}_n , where each subgraph is induced by

certain vertices of \mathcal{G}_n which reside within either a square or a right angled isosceles triangle having side length n_k . The index sets for these partitions are $\{\mathbf{V}_k\}_{k=0}^K$, where $\mathbf{V}_0 = \emptyset$ and $\mathbf{V}_k \subset \mathcal{Z}_{n_0/n_1}^{(a)} \times \dots \times \mathcal{Z}_{n_{k-1}/n_k}^{(a)}$ for $k \in [K]$. For $k = 0$, \mathcal{B}_0 is the singleton trivial partition $\{\mathcal{B}_\emptyset^{(0)} := \mathcal{V}_n^{(a)}\}$. For $k = 1$, let $\mathbf{V}_1 := \{\mathbf{u} \in \mathcal{Z}_{n_0/n_1}^{(a)} : \mathbf{V}_\mathbf{u}^{(1)} \neq \emptyset\}$, where

$$\mathbf{V}_\mathbf{u}^{(1)} := \mathcal{V}_{n_0}^{(a)} \cap [u_1 n_1, u_1 n_1 + n_1) \times [u_2 n_1, u_2 n_1 + n_1).$$

For $\mathbf{u} \in \mathbf{V}_1$, define $\mathcal{B}_\mathbf{u}^{(1)}$ to be the subgraph of \mathcal{G}_n induced by $\mathbf{V}_\mathbf{u}^{(1)}$ and define $\mathcal{B}_1 := \{\mathcal{B}_\mathbf{u}^{(1)} : \mathbf{u} \in \mathbf{V}_1\}$. It is clear that for each $\mathbf{u} \in \mathbf{V}_1$, the vertices belonging to $\mathbf{V}_\mathbf{u}^{(1)}$ reside within either a complete square or a diagonally halved square (right-angled isosceles triangle) having side length n_1 . Also, the vertex sets $\{\mathbf{V}_\mathbf{u}^{(1)} : \mathbf{u} \in \mathbf{V}_1\}$ are disjoint. Having defined \mathcal{B}_k for some $k \in [K]_0$, we obtain \mathcal{B}_{k+1} as follows. Note that $\mathcal{B}_\mathbf{v}^{(k)} \in \mathcal{B}_k$ is the subgraph of \mathcal{G}_n induced by the vertex set $\mathbf{V}_\mathbf{v}^{(k)}$, and the vertices belonging to $\mathbf{V}_\mathbf{v}^{(k)}$ reside within either a complete square or an isosceles triangle having side length n_k . So, $\mathbf{V}_\mathbf{v}^{(k)}$ can be divided into disjoint subsets $\mathbf{V}_{\mathbf{v},\mathbf{u}}^{(k+1)}$, $\mathbf{u} \in \mathcal{Z}_{n_k/n_{k+1}}^{(a)}$ as follows. $\mathbf{V}_\mathbf{v}^{(k)}$ is a spatial translate (say τ) of either $\mathcal{V}_n \cap \{(x, y) : 0 \leq x, y < n_k\}$ or $\mathcal{V}_n \cap \{(x, y) : 0 \leq y \leq x < n_k\}$. We take $\mathbf{V}_{\mathbf{v},\mathbf{u}}^{(k+1)}$ to be the intersection of $\mathbf{V}_\mathbf{v}^{(k)}$ and the image under τ of $[u_1 n_{k+1}, u_1 n_{k+1} + n_{k+1}) \times [u_2 n_{k+1}, u_2 n_{k+1} + n_{k+1})$. Having defined $\mathbf{V}_{\mathbf{v},\mathbf{u}}^{(k+1)}$, we define $\mathbf{V}_{k+1} := \{(\mathbf{v}, \mathbf{u}) : \mathbf{v} \in \mathbf{V}_k, \mathbf{V}_{\mathbf{v},\mathbf{u}}^{(k+1)} \neq \emptyset\}$. For $\mathbf{v} \in \mathbf{V}_{k+1}$, $\mathcal{B}_\mathbf{v}^{(k+1)}$ denotes the subgraph of \mathcal{G}_n induced by vertex subset $\mathbf{V}_\mathbf{v}^{(k+1)}$ and define $\mathcal{B}_{k+1} := \{\mathcal{B}_\mathbf{w}^{(k+1)} : \mathbf{w} \in \mathbf{V}_{k+1}\}$. See Figure 4 for a sketch of these partitions.

After defining the sequence of partitions $\{\mathcal{B}_k\}_{k=0}^K$ as above, we will assign a random variable $Q_\mathbf{v}^{(k)}$ to the subgraph $\mathcal{B}_\mathbf{v}^{(k)}$ for all $\mathbf{v} \in \mathbf{V}_k$ and $0 \leq k \leq K$. These random variables will be defined using backward induction in k . For $\mathbf{v} \in \mathbf{V}_K$, we

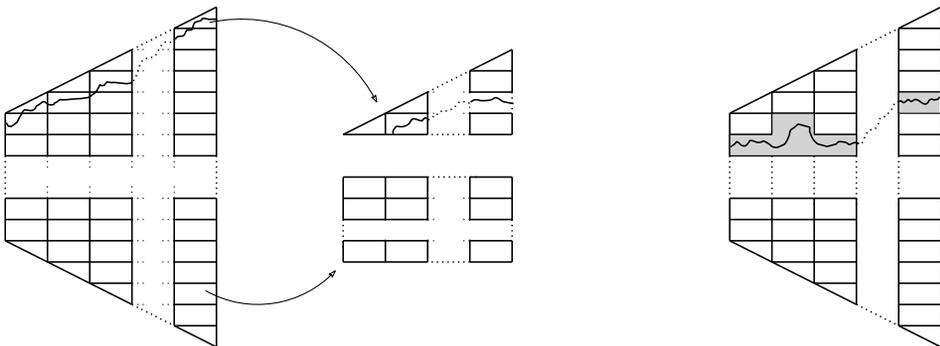


FIG. 4. The left figure gives a sketch of the partitions \mathcal{B}_k , $k = 0, 1, 2$. The whole graph is $\mathcal{B}_\emptyset^{(0)}$. The larger squares correspond to subgraphs in \mathcal{B}_1 and the small squares correspond to subgraphs corresponding to \mathcal{B}_2 . The right figure shows the corresponding generalized path on $G_\emptyset^{(0)}$.

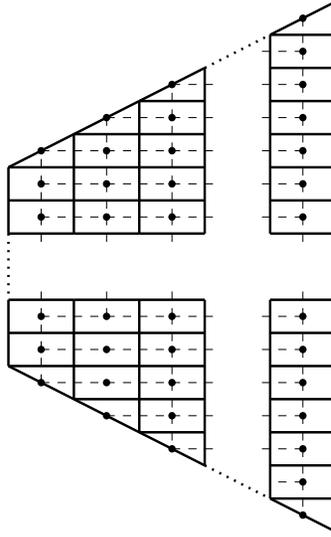


FIG. 5. This is a sketch of the coarse grained graph. Here, the vertex set consists of the squares and the edges are represented by the links.

order the random variables $\{\mathbf{X}_{\mathbf{u}} : \mathbf{u} \in V_{\mathbf{v}}^{(K)}\}$ using the partial order \preceq to obtain the column vector $\mathbf{X}_{\mathbf{v}}^{(K)}$ consisting of $O(n_K^2)$ many entries. Then $Q_{\mathbf{v}}^{(K)}$ is taken to be the quadratic form

$$Q_{\mathbf{v}}^{(K)} := (\mathbf{X}_{\mathbf{v}}^{(K)})^T [\bar{\mathbf{A}}(V_{\mathbf{v}}^{(K)})] \mathbf{X}_{\mathbf{v}}^{(K)},$$

where $[\bar{\mathbf{A}}(\cdot)]$ is as defined in (2.1). Having defined the random variables $\{Q_{\mathbf{v}}^{(l)} : \mathbf{v} \in \mathbf{V}_l\}$ for all $l \in [K] \setminus [k]_0$ we obtain the random variables $\{Q_{\mathbf{u}}^{(k-1)} : \mathbf{u} \in \mathbf{V}_{k-1}\}$ as follows. If $\mathbf{u} \in \mathbf{V}_{k-1}$, then $B_{\mathbf{u}}^{(k-1)}$ can be thought of as a (coarse-grained) graph $\bar{B}_{\mathbf{u}}^{(k-1)}$ (see Figure 5) having the vertex set $\bar{V}_{\mathbf{u}}^{(k-1)} := \{B_{\mathbf{u}, \mathbf{w}}^{(k)} : (\mathbf{u}, \mathbf{w}) \in \mathbf{V}_k\}$ and edge set corresponding to the neighboring relation: $B_{\mathbf{u}, \mathbf{w}}^{(k)} \sim B_{\mathbf{u}, \mathbf{w}'}$ if $\|\mathbf{w} - \mathbf{w}'\|_{\infty} = 1$. For completeness, we define $\bar{B}^{(K)}$ and $\bar{V}^{(K)}$ to be same as $B^{(K)}$ and $V^{(K)}$, respectively. We order the random variables $\{Q_{\mathbf{u}, \mathbf{w}}^{(k)} : (\mathbf{u}, \mathbf{w}) \in \mathbf{V}_k\}$ using the partial order \preceq and obtain the vector $\mathbf{X}_{\mathbf{u}}^{(k-1)}$ consisting of $O((n_{k-1}/n_k)^2)$ entries. Then we define

$$Q_{\mathbf{u}}^{(k-1)} := (\mathbf{X}_{\mathbf{u}}^{(k-1)})^T [\bar{\mathbf{A}}(\bar{V}_{\mathbf{u}}^{(k-1)})] \mathbf{X}_{\mathbf{u}}^{(k-1)},$$

where $[\bar{\mathbf{A}}(\cdot)]$ is as defined in (2.1). We proceed in this way until $Q_{\emptyset}^{(0)}$ is defined. Thus, $Q_{\emptyset}^{(0)}$ is a quadratic form in terms of the random variables $\{Q_{\mathbf{v}}^{(1)} : \mathbf{v} \in \mathbf{V}_1\}$, where each $Q_{\mathbf{v}}^{(1)}$ is again a quadratic form in terms of the random variables $\{Q_{\mathbf{v}, \mathbf{u}}^{(2)} : (\mathbf{v}, \mathbf{u}) \in \mathbf{V}_2\}$, and so on.

Our next goal is to study the distribution of $Q_\emptyset^{(0)}$ under \mathbb{P}_0 and $\mathbb{P}_{1,\pi}$. Choose and fix any path $\pi \in \mathcal{P}(\mathcal{G}_n)$. Note that for each $0 \leq k \leq K$ and $\mathbf{v} \in \mathbf{V}_k$, π induces generalized paths $\Pi_{\mathbf{v}}^{(k)} \in \widetilde{\mathcal{F}}(\bar{B}_{\mathbf{v}}^{(k)})$ for the (possibly coarse-grained) graph $\bar{B}_{\mathbf{v}}^{(k)}$; see Figure 4 for a sketch of an anomalous path π and corresponding (coarse-grained) generalized paths $\Pi_{\emptyset}^{(0)}$ and $\Pi_{\mathbf{v}}^{(1)}$ for some $\mathbf{v} \in \mathbf{V}_1$. Using this fact and Proposition 2.4, we next construct the basic noise variables $(W_{\mathbf{v}}^{(k)}, \mathbf{v} \in \mathbf{V}_k)$, additional noise variables $(U_{\mathbf{v}}^{(k)}, \mathbf{v} \in \mathbf{V}_k)$ and signals $(v_{\mathbf{v}}^{(k)}, \mathbf{v} \in \mathbf{V}_k)$ for $0 \leq k \leq K$ inductively as follows. (Since π is fixed, we eliminate it from the notation, writing, e.g., $U_{\mathbf{v}}^{(K)}$ for $U_{\mathbf{v},\pi}^{(K)}$.)

First, we define the attributes at level K . For $\mathbf{v} \in \mathbf{V}_K$, we apply Proposition 2.4, with n replaced by n_K , basic noise variables $(X_{\mathbf{u}}, \mathbf{u} \in V_{\mathbf{v}}^{(K)})$, which are the basic noise variables for the vertices of \mathcal{G}_n , additional noise variables given by zeros and signals $(\mu \mathbf{1}_{\{\mathbf{u} \in \pi\}}, \mathbf{u} \in V_{\mathbf{v}}^{(K)})$ and obtain the basic noise variable, additional noise variable and signal, which we will denote by $W_{\mathbf{v}}^{(K)}, U_{\mathbf{v}}^{(K)}$ and $v_{\mathbf{v}}^{(K)}$, respectively.

Having obtained $((W_{\mathbf{v}}^{(k)}, U_{\mathbf{v}}^{(k)}, v_{\mathbf{v}}^{(k)}), \mathbf{v} \in \mathbf{V}_k)$ for $k \in [K]$, we obtain the attributes $(W_{\mathbf{v}}^{(k-1)}, U_{\mathbf{v}}^{(k-1)}, v_{\mathbf{v}}^{(k-1)})$ by applying Proposition 2.4 with basic noise variables $(W_{(\mathbf{v},\mathbf{u})}^{(k)}, (\mathbf{v}, \mathbf{u}) \in \bar{V}_{\mathbf{v}}^{(k-1)})$, additional noise variables $(U_{(\mathbf{v},\mathbf{u})}^{(k)}, (\mathbf{v}, \mathbf{u}) \in \bar{V}_{\mathbf{v}}^{(k-1)})$ and signals $(v_{(\mathbf{v},\mathbf{u})}^{(k)}, (\mathbf{v}, \mathbf{u}) \in \bar{V}_{\mathbf{v}}^{(k-1)})$. We proceed in this way until $W_{\emptyset}^{(0)}, U_{\emptyset}^{(0)}$ and $v_{\emptyset}^{(0)}$ are defined.

We next estimate the signals $(v_{\mathbf{v}}^{(k)}, k \in [K], \mathbf{v} \in \mathbf{V}_k)$. To do so, we first define certain vertical segments $(S_{\mathbf{i}}^{(k)}, \mathbf{i} \in \mathbf{I}_k := \otimes_1^k [n_{i-1}/n_i]_0, k \in [K + 1])$ of \mathcal{G}_n , which we call *slabs*. Define

$$S_{\emptyset}^{(0)} := \mathcal{V}_n \quad \text{and}$$

$$S_{\mathbf{i}}^{(k)} := \left\{ \mathbf{v} \in \mathcal{V}_n : \sum_{j=1}^k i_j n_j \leq v_1 < \sum_{j=1}^k i_j n_j + n_k \right\}$$

for $\mathbf{i} \in \mathbf{I}_k$ and $k \in [K + 1]$.

Note that the slabs are line segments if $k = K + 1$. We also define the projection map:

$$p : \bigcup_{k \in [K]} \mathbf{V}_k \mapsto \bigcup_{k \in [K]} \mathbf{I}_k$$

by assigning $p(\mathbf{v}) := \mathbf{i}$, if \mathbf{v}_j has first component i_j for all $j \in [k]$.

We call a subgraph $B_{\mathbf{v}}^{(k)}$ *touched* if π intersects it. We call a subgraph $B_{\mathbf{v}}^{(K)}$ *good* if π intersects it in at least $\frac{1}{2}n_K$ slabs (= line segments) among $S_{(p(\mathbf{v}), i_{K+1})}^{(K+1)}, i_{K+1} \in$

$[n_K]_0$, that is,

$B_{\mathbf{v}}^{(K)}$ is “good”

$$\text{if } \left| \left\{ i_{K+1} \in [n_K]_0 : \pi_i \in B_{\mathbf{v}}^{(K)} \text{ for } i = \sum_{j=1}^K v_{j,1} n_j + i_{K+1} \right\} \right| \geq \frac{1}{2} n_K.$$

We extend the definition of good subgraphs to other levels inductively as follows. For $k \in [K - 1]$, we call a subgraph $B_{\mathbf{v}}^{(k)}$ good if there are at least $\frac{1}{2}(n_k/n_{k+1})$ many slabs among $S_{(\mathbf{p}(\mathbf{v}), i_{k+1})}^{(k+1)}$, $i_{k+1} \in [n_k/n_{k+1}]_0$ where $B_{\mathbf{v}}^{(k)}$ contains at least one good subgraph $B_{(\mathbf{v}, \mathbf{u})}^{(k+1)}$, that is,

$B_{\mathbf{v}}^{(k)}$ is “good”

$$\text{if } \left| \{ i_{k+1} \in [n_k/n_{k+1}]_0 : \exists \text{ good } B_{(\mathbf{v}, \mathbf{u})}^{(k+1)} \text{ satisfying } \mathbf{p}(\mathbf{u}) = i_{k+1} \} \right| \geq \frac{n_k}{2n_{k+1}}.$$

The following lemma contains the required control on the signal variables $v_{\mathbf{v}}^{(k)}$.

LEMMA 2.9. *There is a constant $c \in (0, 1)$ such that the following holds for all $k \in [K + 1]_0$:*

$$(A) \quad v_{\mathbf{v}}^{(k)} \geq \underline{v}^{(k)} := \frac{1}{c} \exp[2^{K-k+1} \log(c\mu)] \prod_{l=k}^K [\log(n_l/n_{l+1})]^{2^{l-k-1}}$$

if $B_{\mathbf{v}}^{(k)}$ is good,

$$(B) \quad v_{\mathbf{v}}^{(k)} \leq \bar{v}^{(k)} := c \exp[2^{K-k+1} \log(\mu/c)] \prod_{l=k}^K [\log(n_l/n_{l+1})]^{2^{l-k-1}}$$

if π intersects $B_{\mathbf{v}}^{(k)}$.

REMARK 2.10. Note that the condition that π intersects $B_{\mathbf{v}}^{(k)}$ in (B) is made only for aesthetic reasons: the claim is obvious otherwise, for then $v_{\mathbf{v}}^{(k)} = 0$.

PROOF OF LEMMA 2.9. We begin with the following facts:

- (I) Each of the slabs $S_{\mathbf{i}}^{(k)}$, $\mathbf{i} \in \mathbf{I}_k$, has at least one good $B_{\mathbf{v}}^{(k)}$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$.
- (II) Each of the slabs $S_{\mathbf{i}}^{(k)}$, $\mathbf{i} \in \mathbf{I}_k$, has at most two touched $B_{\mathbf{v}}^{(k)}$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$.

To see (II), note that since the slabs $S_{\mathbf{i}}^{(k)}$, $\mathbf{i} \in \mathbf{I}_k$, have width n_k , the subgraphs $\{B_{\mathbf{v}}^{(k)} : \mathbf{p}(\mathbf{v}) = \mathbf{i}\}$ constitute a partition of $S_{\mathbf{i}}^{(k)}$, and each $B_{\mathbf{v}}^{(k)}$ resides within a square (or isosceles triangle at the boundary of $S_{\mathbf{i}}^{(k)}$) having side length n_k , it follows that

for each $k \in [K]$, the path π intersects each $S_{\mathbf{i}}^{(k)}$ in either one subgraph $B_{\mathbf{v}}^{(k)} \in \mathcal{B}_k$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$ or in two (consecutive and disjoint) subgraphs $B_{\mathbf{v}}^{(k)}, B_{\mathbf{u}}^{(k)} \in \mathcal{B}_k$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{p}(\mathbf{u}) = \mathbf{i}$. This proves (II).

To see (I), we use induction on k . We first show the induction basis. Since each slab $S_{\mathbf{i}}^{(K)}$ has n_K hyperplanes and π crosses each hyperplane, for each slab $S_{\mathbf{i}}^{(K)}$, there is at least one good subgraph $B_{\mathbf{v}}^{(K)}$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$, showing (I) for $k = K$.

Now suppose (I) hold for $k = l + 1$. So each slab $S_{\mathbf{i}}^{(l+1)}$ has at least one good subgraph $B_{\mathbf{v}}^{(l+1)}$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$, at most two touched subgraphs $B_{\mathbf{v}}^{(l+1)}, B_{\mathbf{u}}^{(l+1)}$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{p}(\mathbf{u}) = \mathbf{i}$, $v_{\mathbf{v}}^{(l+1)} \geq \underline{v}^{(l+1)}$ if $B_{\mathbf{v}}^{(l+1)}$ is good, and $v_{\mathbf{v}}^{(l+1)} \leq \bar{v}^{(l+1)}$ if $B_{\mathbf{v}}^{(l+1)}$ is touched.

Now fix $\mathbf{i} \in \mathbf{I}_l$. $S_{\mathbf{i}}^{(l)}$ consists of n_l/n_{l+1} many slabs of level $l + 1$, namely $S_{\mathbf{i}, i_{l+1}}^{(l+1)}$, $i_{l+1} \in [n_l/n_{l+1}]_0$. Each such slab has at least one good subgraph by Assumption (I) for $k = l + 1$. Also, as mentioned in the beginning of the proof, the portion of π within slab $S_{\mathbf{i}}^{(l)}$ resides in either one subgraph $B_{\mathbf{v}}^{(l)} \in \mathcal{B}_l$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{i}$ or in two (consecutive and disjoint) subgraphs $B_{\mathbf{v}}^{(l)}, B_{\mathbf{u}}^{(l)} \in \mathcal{B}_l$ satisfying $\mathbf{p}(\mathbf{v}) = \mathbf{p}(\mathbf{u}) = \mathbf{i}$. In the first case, it is obvious that $B_{\mathbf{v}}^{(l)}$ is good. In the second case, if both $B_{\mathbf{v}}^{(l)}$ and $B_{\mathbf{u}}^{(l)}$ are not good, then there will be at least one slab among $S_{\mathbf{i}, i_{l+1}}^{(l+1)}$, $i_{l+1} \in [n_l/n_{l+1}]_0$, having no good subgraph of level $l + 1$ within it. This leads to a contradiction to the induction hypothesis concerning (I). We conclude that (I) holds for all $k \in [K + 1]_0$.

We now turn to the proof of (A), (B). Again, the proof is by backward induction on k . We begin with proving the basis of the induction. Suppose that $B_{\mathbf{v}}^{(K)}$ is good. Then, using Proposition 2.4, particularly the lower bound in (2.6), we get that $v_{\mathbf{v}}^{(K)} \geq c\mu^2\sqrt{\log(n_K)}$, thus showing (A) in the case $k = K$. On the other hand, the upper bound in (2.6) implies $v_{\mathbf{v}}^{(K)} \leq c^{-1}\mu^2\sqrt{\log(n_K)}$ whenever π intersects $B_{\mathbf{v}}^{(K)}$, showing (B) in the case $k = K$. This completes the proof of the base of the induction.

Now suppose (A), (B) hold for $k = l + 1$. Suppose $B_{\mathbf{v}}^{(l)}$ is good. Then, using Proposition 2.4, particularly the lower bound in (2.6), and noting that $\bar{B}_{\mathbf{v}}^{(l)}$ resides within either a square or an isosceles triangle having side length n_l/n_{l+1} , $v_{\mathbf{v}}^{(l)} \geq c(\underline{v}^{(l+1)})^2\sqrt{\log(n_l/n_{l+1})}$. On the other hand, the upper bound in (2.6) implies $v_{\mathbf{v}}^{(l)} \leq \frac{1}{c}(\bar{v}^{(l+1)})^2\sqrt{\log(n_l/n_{l+1})}$ whenever π intersects $B_{\mathbf{v}}^{(l)}$. Combining these with the expressions of $\underline{v}^{(l+1)}$ and $\bar{v}^{(l+1)}$ [obtained from (2.10)] we see that (A) and (B) of (2.10) hold for $k = l$. Thus, all assertions of (2.10) are true for $k = l$, and the induction argument is complete. \square

We return to the proof of the upper bound of Theorem 1.1. Note first that the choice of constants K and ε_k made in (2.13) and (2.16) below, together with Lemma 2.9(B), ensure that $\max_{\mathbf{u} \in \mathbf{V}_k} v_{\mathbf{u}}^{(k)} \leq 1$. By construction it follows, using

Proposition 2.4, that for all $0 \leq k \leq K$,

$$Q_{\mathbf{v}}^{(k)} \stackrel{d}{=} \begin{cases} W_{\mathbf{v}}^{(k)} & \text{under } H_0, \\ W_{\mathbf{v}}^{(k)} + U_{\mathbf{v}}^{(k)} + \nu_{\mathbf{v}}^{(k)} & \text{under } H_{1,\pi}, \end{cases}$$

$$\mathbb{E}U_{\mathbf{v}}^{(k)} = 0, \quad \mathbb{E}[U_{\mathbf{v}}^{(k)}]^2 \leq C \frac{1}{\log(n_k/n_{k+1})} \max_{\mathbf{u}: (\mathbf{v}, \mathbf{u}) \in \mathbf{V}_{k+1}} (\nu_{(\mathbf{v}, \mathbf{u})}^{(k+1)})^2.$$

In particular, $\mathbb{P}_0(Q_{\emptyset}^{(0)} \leq \cdot) = \mathbb{P}(W_{\emptyset}^{(0)} \leq \cdot)$, so using (2.5) with n replaced by $|\mathbf{V}_1| \asymp (n_0/n_1)^2 = n^{2-2\varepsilon_1}$,

$$\sup_{x \in \mathbb{R}} |\mathbb{P}_0(Q_{\emptyset}^{(0)} \leq x) - \Phi(x)| \leq C \left[(\log n^{2-2\varepsilon_1})^{-1/2} + \left(\frac{\beta_3^{(1)}}{n^{2-2\varepsilon_1}} + \frac{(\beta_3^{(1)})^2}{n^{2-2\varepsilon_1} \log^{3/2} n^{2-2\varepsilon_1}} \right)^{1/4} \right],$$

where $\beta_3^{(1)} := \max_{\mathbf{v} \in \mathbf{V}_1} \mathbb{E}(|W_{\mathbf{v}}^{(1)}|^3)$. In order to bound $\beta_3^{(1)}$, we will use Theorem 2.5. If we define $\beta_s^{(k)} := \max_{\mathbf{v} \in \mathbf{V}_k} \mathbb{E}(|W_{\mathbf{v}}^{(k)}|^s)$, then Theorem 2.5 gives $\beta_s^{(k)} \leq C_s \beta_{2s}^{(k+1)}$ for all $k \in [K - 1]$, where $C_s = 2^{5s/2} \Gamma((s + 1)/2) \Gamma(s + 1/2)^{1/2}$. Also $\beta_s^{(K)} \asymp \int_{\mathbb{R}} |x|^s d\Phi(x) \asymp \Gamma((s + 1)/2)$. Combining the last two observations,

$$\beta_3^{(1)} \leq \exp \left[\frac{5}{2} \log 2 \sum_{l=0}^{K-1} 3 \cdot 2^l + \sum_{l=0}^K \log \Gamma \left(\frac{1}{2} + \frac{3 \cdot 2^l}{2} \right) + \frac{1}{2} \sum_{l=0}^{K-1} \log \Gamma \left(\frac{1}{2} + 3 \cdot 2^l \right) \right] \leq \exp(C K^2 2^K)$$

for some $C > 0$. The last inequality holds because $\Gamma(k) \asymp k^k$. Combining the last two displays, we get

$$(2.11) \quad \sup_{x \in \mathbb{R}} |\mathbb{P}_0(Q_{\emptyset}^{(0)} \leq x) - \Phi(x)| \leq C \left[(\log n^{2-2\varepsilon_1})^{-1/2} + \exp \left(C K^2 2^K - \frac{1}{2} (1 - \varepsilon_1) \log n \right) \right]$$

for some constant $C > 0$.

Next, we simplify the formula for $\underline{\nu}^{(0)}, \bar{\nu}^{(0)}$, which are defined in Lemma 2.9, and then specify the constants $(\varepsilon_k, k \in [K])$. Clearly,

$$\underline{\nu}^{(0)} = \frac{1}{c} \exp[2^{K+1} \log(c\mu)] \prod_{l=0}^K [\log(n_l/n_{l+1})]^{2^{l-1}} \quad \text{and}$$

$$\bar{\nu}^{(0)} = c \exp[2^{K+1} \log(\mu/c)] \prod_{l=0}^K [\log(n_l/n_{l+1})]^{2^{l-1}}.$$

Recalling that $n_0 = n$, $\varepsilon_{K+1} = 0$ and $n_k = n_{k-1}^{\varepsilon_k}$, $k \in [K]$, it is easy to see that

$$\begin{aligned}
 \underline{v}^{(0)} &= \frac{1}{c\sqrt{\log n}} \exp[2^{K+1} \log(c\mu\sqrt{\log n})] \\
 &\quad \times \prod_{l=0}^K [\varepsilon_1 \varepsilon_2 \cdots \varepsilon_l (1 - \varepsilon_{l+1})]^{2^{l-1}} \quad \text{and} \\
 \bar{v}^{(0)} &= \frac{c}{\sqrt{\log n}} \exp[2^{K+1} \log(\mu\sqrt{\log n/c})] \\
 &\quad \times \prod_{l=0}^K [\varepsilon_1 \varepsilon_2 \cdots \varepsilon_l (1 - \varepsilon_{l+1})]^{2^{l-1}}.
 \end{aligned}
 \tag{2.12}$$

Rearranging the terms, the product in the above display equals

$$\prod_{s=1}^K (1 - \varepsilon_s)^{2^{s-2}} \varepsilon_s^{2^K - 2^{s-1}} = \prod_{s=1}^K [(1 - \varepsilon_s) \varepsilon_s^{2^{K-s+2} - 2}]^{2^{s-2}}.$$

It is not difficult to see that the above product will be maximized if $1 - \varepsilon_s = (2^{K-s+2} - 1)^{-1}$, $s \in [K]$.

$$\tag{2.13} \quad \text{We take } \varepsilon_s = 1 - 2^{-(K-s+2)} \quad \text{for all } s \in [K].$$

With this choice of $(\varepsilon_k, k \in [K])$, and using the fact that $1 - x \leq e^{-x}$ for any x , we see that

$$\begin{aligned}
 &\log \prod_{s=1}^K [(1 - \varepsilon_s) \varepsilon_s^{2^{K-s+2} - 2}]^{2^{s-2}} \\
 &\leq -\log 2 \sum_{s=1}^K (K - s + 2) 2^{s-2} - \sum_{s=1}^K 2^{s-2} (1 - 2^{-(K-s+1)}) \\
 &\leq -\log 2 \sum_{s=2}^{K+1} s 2^{K-s} - \frac{1}{2} \sum_{s=1}^K 2^{s-2} \leq -\underline{C} 2^{K+1}
 \end{aligned}$$

for some constant $\underline{C} > 0$. For the same choice of $(\varepsilon_k, k \in [K])$, using the inequality $(1 - \varepsilon)^k \geq 1 - k\varepsilon$ for any $k > 0$, we get

$$\begin{aligned}
 &\log \prod_{s=1}^K [(1 - \varepsilon_s) \varepsilon_s^{2^{K-s+2} - 2}]^{2^{s-2}} \\
 &\geq -\log 2 \left[\sum_{s=1}^K (K - s + 2) 2^{s-2} - \sum_{s=1}^K (1 - (K - s + 2)) 2^{s-2} \right] \\
 &= -\log 2 \left[\sum_{s=2}^{K+1} s 2^{K-s+1} - \sum_{s=1}^K 2^{s-2} \right] \geq -\bar{C} 2^{K+1}
 \end{aligned}$$

for some constant $\bar{C} > \underline{C}$. Combining the last two displays with (2.12),

$$(2.14) \quad \begin{aligned} v_{\varnothing}^{(0)} \geq \underline{v}^{(0)} &\geq \frac{1}{c\sqrt{\log n}} \exp[2^{K+1}(\log(c\mu\sqrt{\log n}) - \bar{C})] \quad \text{and} \\ v_{\varnothing}^{(0)} \leq \bar{v}^{(0)} &\leq \frac{c}{\sqrt{\log n}} \exp[2^{K+1}(\log(\mu\sqrt{\log n}/c) - \underline{C})]. \end{aligned}$$

Applying (2) of Proposition 2.4, and using (a) the above bound for $v_{\varnothing}^{(0)}$, (b) the value of ε_1 [as specified in (2.13)] and (c) the relation between $\bar{v}^{(0)}$ and $\bar{v}^{(1)}$ (as defined in Lemma 2.9), we see that there is a constant $C_2 > 0$ such that

$$(2.15) \quad \begin{aligned} \mathbb{E}[(U_{\varnothing}^{(0)})^2] &\leq C_2 \frac{(\bar{v}^{(1)})^2}{\log(n/n_1)} = C_2 c \frac{\bar{v}^{(0)}}{\log^{3/2}(n/n_1)} \\ &\leq C_2 c^2 \frac{2^{\frac{3}{2}(K+1)}}{\log^2 n} \exp\left[2^{K+1}\left(\log\left(\frac{1}{c}\mu\sqrt{\log n}\right) - \underline{C}\right)\right]. \end{aligned}$$

Finally, we specify the value of K . Since $\bar{C} > \underline{C} > 0$ and $c \in (0, 1)$, one can choose $\Delta_0 > 0$ large enough so that for any $\Delta \geq \Delta_0$, $\log(c\Delta) - \bar{C} > 0$ and $\log(\frac{1}{c}\Delta) - \underline{C} \leq \frac{5}{4}(\log(c\Delta) - \bar{C})$. For such a Δ_0 , if we assume that $\mu\sqrt{\log n} \geq \Delta_0$, and choose K so that

$$(2.16) \quad 2^{K+1}[\log(c\mu\sqrt{\log n}) - \bar{C}] = \log \log n,$$

then $2^{K+1} \leq (\log \log n)/(\log(c\Delta_0) - \bar{C})$, which ensures $\mathbb{E}[(U_{\varnothing}^{(0)})^2] \leq C_3(\log \log n)^{3/2}(\log n)^{-3/4}$ [using (2.15)] for some constant C_3 and the upper bound in (2.11) is $o(1)$, and $v_{\varnothing}^{(0)} \geq \sqrt{\log n}/c$ [using (2.14)]. So for this choice of K , $W_{\varnothing}^{(0)} \xrightarrow{d} N(0, 1)$, $U_{\varnothing}^{(0)} \xrightarrow{P} 0$ and $v_{\varnothing}^{(0)} \rightarrow \infty$. So if one rejects the null hypothesis when $Q_{\varnothing}^{(0)}$ exceeds $v_{\varnothing}^{(0)}/2$, then its minimax risk will be $o(1)$. \square

2.2. Proof of the lower bound of detection threshold.

PROOF OF THEOREM 1.1, LOWER BOUND. Since the hypothesis testing problem $(\mathcal{P}(\mathcal{G}_n^{(a)}), \mu_n, \Phi)$ has a signal hypothesis which is strictly larger than that of $(\mathcal{P}(\mathcal{G}_n^{(0)}), \mu_n, \Phi)$, the asymptotically powerless part of Theorem 1.1 follows from [1], Theorem 1.1. \square

3. Proof of the supporting results.

3.1. Proof of Lemma 2.1.

PROOF OF LEMMA 2.1. (1) Note that

$$\|[\mathbf{A}(\mathcal{Y}_n)]\|_F^2 = 2 \sum_{(i,u) \in \mathcal{Y}_n} \sum_{(j,v): (i,u) \prec (j,v)} (j-i)^{-2}.$$

In order to estimate the above sum, note that for fixed i, u, j the number of choices of v satisfying $(i, u) \lesssim (j, v)$ is $\asymp |j - i|$. Combining this with the fact that $\sum_{1 \leq k \leq i} k^{-1} \asymp \log i$ implies that

$$\begin{aligned}
 \|\mathbf{A}(\mathcal{Y}_n)\|_F^2 &\asymp \sum_{i,u,j:(i,u) \in \mathcal{Y}_n \text{ and } j > i} (j - i)^{-1} \\
 (3.1) \qquad \qquad \qquad &\asymp \sum_{i,u:(i,u) \in \mathcal{Y}_n} \log(n - i) \asymp n^2 \log n.
 \end{aligned}$$

(2) In order to bound the spectral norm of $[\mathbf{A}(\mathcal{Y}_n)]$, we will use the well-known fact that for any matrix $\mathbf{B} \in \mathbb{R}^{k \times l}$,

$$\max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{B}\} \leq \max_{i \in [k]} \sum_{j \in [l]} |b_{i,j}|.$$

In our case, since $[\mathbf{A}(\mathcal{Y}_n)]$ is symmetric, $\|[\mathbf{A}(\mathcal{Y}_n)]\|$ equals the largest absolute eigenvalue of $[\mathbf{A}(\mathcal{Y}_n)]$, so $\|[\mathbf{A}(\mathcal{Y}_n)]\| \leq \max_{(i,u) \in \mathcal{Y}_n} \sum_{(j,v) \in \mathcal{Y}_n} [A(\mathcal{Y}_n)]_{(i,u),(j,v)}$. Note that for each $(i, u) \in \mathcal{Y}_n$,

$$\sum_{(j,v) \in \mathcal{Y}_n} [A(\mathcal{Y}_n)]_{(i,u),(j,v)} = \sum_{j \in [n], j \neq i} \sum_{v:(j,v) \in [(i,u)]} \frac{1}{|i - j|} \asymp \sum_{j \in [n], j \neq i} 1 \asymp n,$$

which gives the result.

(3) For any $\pi \in \mathcal{P}_n$ and $i \in [n]_0$, let $\pi_i \in \mathbb{Z}$ be such that $(i, \pi_i) \in \pi$. Then

$$\begin{aligned}
 \mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)] \mathbb{1}_\pi &= 2 \sum_{0 \leq i < j < n} [\mathbf{A}(\mathcal{Y}_n)]_{(i,\pi_i),(j,\pi_j)} \\
 &= 2 \sum_{0 \leq i < j < n} (j - i)^{-1} \asymp \sum_{i \in [m]_0} \log(n - i) \asymp n \log n.
 \end{aligned}$$

(4) For $\pi \in \mathcal{P}_n$, $\mathbb{1}_\pi^T [\mathbf{A}(\mathcal{Y}_n)]^2 \mathbb{1}_\pi$ equals

$$\begin{aligned}
 \sum_{i,j \in [n]_0} ([\mathbf{A}(\mathcal{Y}_n)]^2)_{(i,\pi_i),(j,\pi_j)} &= 2 \sum_{0 \leq i < j < n} \sum_{(k,v) \in [(i,\pi_i)] \cap [(j,\pi_j)]} (|i - k| \cdot |j - k|)^{-1} \\
 &\quad + \sum_{i \in [n]_0} \sum_{(k,v) \in [(i,\pi_i)]} (k - i)^{-2} =: I_1 + I_2.
 \end{aligned}$$

In order to estimate I_2 , note that for any $i, k \in [n]_0$, $|\{v : (k, v) \leftrightarrow (i, \pi_i)\}| \asymp |k - i|$. So

$$(3.2) \qquad \qquad \qquad I_2 \asymp \sum_{i \in [n]_0} \sum_{k \in [n]_0, k \neq i} |k - i|^{-1} \asymp n \log n.$$

In order to estimate I_1 , note that the (i, j) th inner sum of I_1 equals

$$\begin{aligned}
 & \left[\sum_{k=0}^{i-1} \sum_{v:(k,v) \rightsquigarrow (i,\pi_i), (j,\pi_j)} (i-k)^{-1}(j-k)^{-1} \right. \\
 & + \sum_{k=j+1}^{n-1} \sum_{v:(k,v) \rightsquigarrow (i,\pi_i), (j,\pi_j)} (k-i)^{-1}(k-j)^{-1} \\
 & \left. + \sum_{k:i < k < j} \sum_{v:(k,v) \rightsquigarrow (i,\pi_i), (j,\pi_j)} (k-i)^{-1}(j-k)^{-1} \right] \\
 & =: I_{1,1}^{i,j} + I_{1,2}^{i,j} + I_{1,3}^{i,j}.
 \end{aligned}
 \tag{3.3}$$

For the sum $I_{1,1}^{i,j}$, we see that $(k, v) \rightsquigarrow (i, \pi_i)$ implies $(k, v) \rightsquigarrow (j, \pi_j)$. So the number of summands in the k th inner sum is $\asymp (i - k)$. Hence,

$$I_{1,1}^{i,j} \asymp \sum_{k=0}^{i-1} (j-k)^{-1} \asymp \int_{j-i}^j \frac{1}{x} dx = \log \frac{j}{j-i},
 \tag{3.4}$$

and so

$$\begin{aligned}
 \sum_{0 \leq i < j < n} I_{1,1}^{i,j} & \asymp \sum_{0 \leq i < j < n} \log j - \sum_{0 \leq i < j < n} \log(j-i) \\
 & = \sum_{j=1}^{n-1} j \log j - \sum_{i=0}^{n-2} \sum_{k=1}^{n-1-i} \log k \\
 & = \sum_{j=1}^{n-1} j \log j - \sum_{k=1}^{n-1} (n-k) \log k \\
 & = \sum_{j=1}^{n-1} 2j \log j - n \sum_{j=1}^{n-1} \log j \\
 & \asymp \int_1^n 2x \log x dx - n \int_1^n \log x dx \\
 & = n^2 \log n - (n^2 - 1)/2 - n^2 \log n + n(n-1) \asymp n^2.
 \end{aligned}$$

Using a similar argument, $I_{1,2}^{i,j} \asymp \sum_{k=j+1}^{n-1} (k-i)^{-1} \asymp \log \frac{n-i-1}{j}$, so

$$\sum_{0 \leq i < j < n} I_{1,2}^{i,j} \asymp n^2.$$

For the sum $I_{1,3}^{i,j}$, the number of summands in the inner sum is at most $2(k - i)$ [resp. $2(j - k)$] when $k < (i + j)/2$ [resp. $k \geq (i + j)/2$]. Thus

$$(3.5) \quad \begin{aligned} I_{1,3}^{i,j} &\leq \sum_{k:i+1 \leq k \leq (i+j)/2} (j-k)^{-1} + \sum_{k:(i+j)/2 < k < j} (k-i)^{-1} \\ &\asymp \int_{(j-i)/2}^{j-i} \frac{1}{x} dx \asymp 1 \quad \text{so} \quad \sum_{0 \leq i < j < n} I_{1,3}^{i,j} \leq Cn^2 \end{aligned}$$

for some constant C . Combing the last three displays, $I_1 = \sum_{k=1}^3 \sum_{0 \leq i < j < n} I_{1,k}^{i,j} \asymp n^2$. This together with (3.2) gives the result.

(5) For $\pi \in \mathcal{P}_n$, $\mathbb{1}_\pi^T [\mathbf{A}(\mathcal{V}_n)] \text{Diag}(\mathbb{1}_\pi) [\mathbf{A}(\mathcal{V}_n)] \mathbb{1}_\pi$ equals

$$\begin{aligned} &\sum_{i,j \in [n]_0} ([\mathbf{A}(\mathcal{V}_n)] \text{Diag}(\mathbb{1}_\pi) [\mathbf{A}(\mathcal{V}_n)])_{(i,\pi_i),(j,\pi_j)} \\ &= 2 \sum_{0 \leq i < j < n} \sum_{k \in [n]_0, k \neq i,j} (|i-k| \cdot |j-k|)^{-1} \\ &\quad + \sum_{i \in [n]_0} \sum_{k \in [n]_0, k \neq i} (k-i)^{-2} =: J_1 + J_2. \end{aligned}$$

In order to estimate J_2 note that for any $i \in [n]_0$, $\sum_{k \in [n]_0, k \neq i} (k-i)^{-2} \asymp C$, so $J_2 \asymp n$. In order to estimate J_1 , note that the (i, j) th inner sum of J_1 equals

$$\begin{aligned} &\left[\sum_{k=0}^{i-1} (i-k)^{-1} (j-k)^{-1} + \sum_{k=j+1}^{n-1} (k-i)^{-1} (k-j)^{-1} \right. \\ &\quad \left. + \sum_{k:i < k < j} (k-i)^{-1} (j-k)^{-1} \right] =: J_{1,1}^{i,j} + J_{1,2}^{i,j} + J_{1,3}^{i,j}. \end{aligned}$$

For the sum $J_{1,1}^{i,j}$, we see that

$$\begin{aligned} J_{1,1}^{i,j} &= \sum_{k=0}^{i-1} \frac{1}{j-i} [(i-k)^{-1} - (j-k)^{-1}] \asymp \frac{1}{j-i} \left[\int_1^i \frac{1}{x} dx - \int_{j-i}^j \frac{1}{x} dx \right] \\ &= \frac{1}{j-i} \log \frac{i(j-i)}{j}, \end{aligned}$$

and so

$$\begin{aligned} \sum_{0 \leq i < j < n} J_{1,1}^{i,j} &\asymp \sum_{0 \leq i < n} \log i \sum_{j:i < j < n} \frac{1}{j-i} - \sum_{0 < j < n} \log j \sum_{0 \leq i < j} \frac{1}{j-i} \\ &\quad + \sum_{0 \leq i < n} \sum_{j-i=1}^{n-i} \frac{\log(j-i)}{j-i} \end{aligned}$$

$$\begin{aligned} &\asymp \sum_{0 \leq i < n} \log i \log(n-i) - \sum_{0 < j < n} (\log j)^2 + \sum_{0 \leq i < n} \frac{1}{2} (\log(n-i))^2 \\ &\leq \log n \int_1^n \log x \, dx - \frac{1}{2} \int_1^n (\log x)^2 \, dx = \frac{1}{2} n (\log n)^2. \end{aligned}$$

Using a similar argument,

$$J_{1,2}^{i,j} \asymp \sum_{k=j+1}^{n-1} \frac{1}{j-i} [(k-j)^{-1} - (k-i)^{-1}] \asymp \frac{1}{j-i} \log \frac{(n-j)(j-i)}{n-i-1}$$

and, therefore, $\sum_{0 \leq i < j < n} J_{1,2}^{i,j} \asymp n(\log n)^2/2$. For the sum $J_{1,3}^{i,j}$, we have that

$$\begin{aligned} J_{1,3}^{i,j} &\leq \sum_{k:i+1 \leq k \leq j-1} \frac{1}{j-i} [(k-i)^{-1} + (j-k)^{-1}] \\ &\asymp \frac{1}{j-i} \int_1^{j-i} \frac{1}{x} \, dx \asymp \frac{\log(j-i)}{j-i}, \end{aligned}$$

and, therefore,

$$\begin{aligned} \sum_{0 \leq i < j < n} J_{1,3}^{i,j} &= \sum_{0 \leq i < n} \sum_{j=i+1}^{n-1} \frac{\log(j-i)}{j-i} \asymp \sum_{0 \leq i < n} \int_1^{n-i} \frac{\log x}{x} \, dx \\ &\asymp \sum_{0 \leq i < n} (\log(n-i))^2 \asymp n(\log n)^2. \end{aligned}$$

Combing the last estimates, we obtain that $J_1 = \sum_{k=1}^3 \sum_{0 \leq i < j < n} J_{1,k}^{i,j} \asymp n(\log n)^2$. This together with the bound on J_2 gives the result.

(6) Note that $\|[\mathbf{A}(\mathcal{Y}_n)]^2\|_F^2$ equals

$$\begin{aligned} &\sum_{\mathbf{u}, \mathbf{v} \in \mathcal{Y}_n} (([\mathbf{A}(\mathcal{Y}_n)]^2)_{\mathbf{u}, \mathbf{v}})^2 \\ &\asymp n^2 \left[\sum_{0 \leq u_1 < v_1 < n} \left(\sum_{\mathbf{w} \in [\mathbf{u}] \cap [\mathbf{v}]} (|u_1 - w_1| \cdot |v_1 - w_1|)^{-1} \right)^2 \right. \\ &\quad \left. + \sum_{u_1 \in [n]_0} \left(\sum_{\mathbf{w} \in [\mathbf{u}]} (w_1 - u_1)^{-2} \right)^2 \right] =: K_1 + K_2. \end{aligned}$$

In order to estimate K_2 , note that for any $u_1, w_1 \in [n]_0$, $|\{w_2 : \mathbf{w} \leftrightarrow \mathbf{u}\}| \asymp |w_1 - u_1|$. So

$$K_2 \asymp n^2 \sum_{u_1 \in [n]_0} \left(\sum_{w_1 \in [n]_0, w_1 \neq u_1} |w_1 - u_1|^{-1} \right)^2 \asymp n^3 \log^2 n.$$

Next, note that the (u_1, v_1) th inner sum of K_1 equals n^2 times

$$(I_{1,1}^{u_1, v_1} + I_{1,2}^{u_1, v_1} + I_{1,3}^{u_1, v_1})^2 \asymp [(I_{1,1}^{u_1, v_1})^2 + (I_{1,2}^{u_1, v_1})^2 + (I_{1,3}^{u_1, v_1})^2],$$

where $I_{1,i}^{u_1, v_1}$, $i = 1, 2, 3$, are as in (3.3). We have that $I_{1,1}^{u_1, v_1} \asymp \log \frac{v_1}{v_1 - u_1}$ [see (3.4)] and, therefore,

$$\begin{aligned} & \sum_{0 \leq u_1 < v_1 < n} (I_{1,1}^{u_1, v_1})^2 \\ & \asymp \sum_{0 \leq u_1 < v_1 < n} \log^2 v_1 - \sum_{0 \leq u_1 < v_1 < n} \log^2 (v_1 - u_1) \\ & = \sum_{j=1}^{n-1} j \log^2 j - \sum_{i=0}^{n-2} \sum_{k=1}^{n-1-i} \log^2 k = \sum_{j=1}^{n-1} j \log^2 j - \sum_{k=1}^{n-1} (n-k) \log^2 k \\ & = \sum_{j=1}^{n-1} 2j \log^2 j - n \sum_{j=1}^{n-1} \log^2 j \asymp \int_1^n 2x \log^2 x \, dx - n \int_1^n \log^2 x \, dx \\ & = n^2(\log^2 n - \log n) - \frac{n^2 - 1}{2} - n^2(\log^2 n - 2 \log n) + 2n(n - 1) \\ & \asymp n^2 \log n. \end{aligned}$$

Using similar argument, $\sum_{0 \leq i < j < n} (I_{1,2}^{i,j})^2 \asymp n^2 \log n$. We also have that

$$\sum_{0 \leq u_1 < v_1 < n} (I_{1,3}^{u_1, v_1})^2 \leq Cn^2;$$

see (3.5). Combing the last estimates, I_1 and hence $\|[\mathbf{A}(\mathcal{Y}_n)]\|_F^2$ is $\asymp n^4 \log n$. This together with Lemma 2.1(1) gives the result. \square

3.2. Proof of Proposition 2.8.

PROOF OF PROPOSITION 2.8. Using Theorem 2.7 and Lemma 2.1(2),

$$(3.6) \quad \sup_{x \in \mathbb{R}} |\mathbb{P}(G_n([\bar{\mathbf{A}}(\mathcal{Y}_n)]) \leq x) - \Phi(x)| \leq C(\log n)^{-1/2}.$$

We obtain $\nu_{\mathbf{v}}$, $s_{\mathbf{v}}^2$, Δ and L [as defined in (2.9)] for $[\bar{\mathbf{A}}(\mathcal{Y}_n)]$ and $(U_{\mathbf{v}}, \mathbf{v} \in \mathcal{Y}_n)$. By Lemma 2.1(6), we have $\Delta \asymp (\log n)^{-1/2}$, so we can apply Theorem 2.6. Suppose $F_{\mathbf{v}}$ denotes the distribution function of $U_{\mathbf{v}}$ and let $\bar{F}_{\mathbf{v}} := 1 - F_{\mathbf{v}}$ and $\bar{\Phi} := 1 - \Phi$. Then

$$\begin{aligned} \nu_{\mathbf{v}} & \leq \int_0^\infty 3x^2(\bar{F}_{\mathbf{v}}(x) + \bar{\Phi}(x)) \, dx + \int_{-\infty}^0 3x^2(F_{\mathbf{v}}(x) + \Phi(x)) \, dx \\ & = \int_0^\infty (\bar{F}_{\mathbf{v}}(x) + \bar{\Phi}(x) + F_{\mathbf{v}}(-x) + \Phi(-x)) \, d(x^3) \end{aligned}$$

$$\begin{aligned}
 &= \int_0^\infty (\mathbb{P}(|X_{F_v}|^3 > y) + \mathbb{P}(|X_\Phi|^3 > y)) dy \\
 &= \int_{\mathbb{R}} |x|^3 d(F_v + \Phi)(x) \leq \beta_3 + 4/\sqrt{2\pi}.
 \end{aligned}$$

We have that $\sum_{\mathbf{u} \in \mathcal{Y}_n} [\mathbf{A}(\mathcal{Y}_n)]_{\mathbf{v}, \mathbf{u}}^2 \asymp \log n$ and $\sum_{\mathbf{u} \in \mathcal{Y}_n} [\mathbf{A}(\mathcal{Y}_n)]_{\mathbf{v}, \mathbf{u}}^3 \asymp 1$ for each $\mathbf{v} \in \mathcal{Y}_n$, by the argument leading to (3.1). This together with Lemma 2.1(1) gives $s_{\mathbf{v}}^2 \asymp n^{-2}$, $\sum_{\mathbf{u}, \mathbf{v} \in \mathcal{Y}_n} [\bar{\mathbf{A}}(\mathcal{Y}_n)]_{\mathbf{v}, \mathbf{u}}^3 \asymp \frac{1}{n \log^{3/2} n}$. These estimates give $L \leq \frac{C}{n} (1 + \beta_3 + \frac{\beta_3^2}{\log^{3/2} n})$. Using that, Theorem 2.6 and (3.6) give the claimed bound. \square

3.3. *Proof of Proposition 2.4.* We begin with the following observation.

LEMMA 3.1. $\min_{I \subset [n]} \max_{S \in \{I, [n] \setminus I\}} \sum_{i, j \in S, i \neq j} |j - i|^{-1} \geq cn \log n$ for some constant $c > 0$.

PROOF. Let $A_n := \min_{I \subset [n]} \max_{S \in \{I, [n] \setminus I\}} \sum_{i, j \in S, i \neq j} |j - i|^{-1}$. We have that

$$\begin{aligned}
 (3.7) \quad A_n &\geq \frac{1}{2} \min_{I \subset [n]} \left(\sum_{i, j \in I, i \neq j} |j - i|^{-1} + \sum_{i, j \in [n] \setminus I, i \neq j} |j - i|^{-1} \right) \\
 &\geq \frac{1}{2} \min_{I \subset [n], |I| \geq n/2} \sum_{i, j \in I, i \neq j} |j - i|^{-1} \\
 &= \frac{1}{2} \min_{\alpha \in [1/2, 1]} \min_{I \subset [n], |I| = \lfloor \alpha n \rfloor} \sum_{i, j \in I, i \neq j} |j - i|^{-1} \\
 &=: \frac{1}{2} \min_{\alpha \in [1/2, 1]} B_n(\alpha).
 \end{aligned}$$

Since $\alpha \mapsto B_n(\alpha)$ is monotone, the lemma will follow from (3.7) if we show the existence of a constant $c > 0$ so that, for all n integer,

$$(3.8) \quad B_n(1/2) \geq cn \log n.$$

To prove (3.8), we begin by claiming that there exists $\alpha_0 < 1$ and a constant $c > 0$ so that

$$(3.9) \quad B_n(\alpha_0) > cn \log n.$$

Indeed, with $|I| = \alpha n$,

$$\begin{aligned}
 \sum_{i, j \in I, i \neq j} |j - i|^{-1} &= \sum_{i, j \in [n], i \neq j} |j - i|^{-1} - \sum_{i, j \in [n] \setminus I, i \neq j} |j - i|^{-1} \\
 &\quad - 2 \sum_{i \in [n] \setminus I, j \in I} |j - i|^{-1}
 \end{aligned}$$

$$\begin{aligned} &\geq n \log n(1 + o(1)) - 2 \sum_{i \in [n] \setminus I, j \in [n], j \neq i} |j - i|^{-1} \\ &= (1 - 4(1 - \alpha))n \log n(1 + o(1)). \end{aligned}$$

In particular, (3.9) holds with $\alpha_0 = 4/5$ and $c = 1/5$. Note that unfortunately, we cannot yet take $\alpha_0 = 1/2$.

Continuing with the proof of (3.8), note that by rescaling, for any $\alpha \in (0, 1)$,

$$(3.10) \quad B_{n/2}(\alpha) = 2B_n(\alpha/2).$$

Thus, with $\beta_0 := \alpha_0/2 < 1/2$,

$$B_n(\beta_0) = \frac{1}{2}B_{n/2}(2\beta_0) = \frac{1}{2}B_{n/2}(\alpha_0) \geq \frac{c}{4}n \log n.$$

By the monotonicity of $\alpha \mapsto B_n(\alpha)$, this proves (3.8), and hence the lemma. \square

Combining Lemma 3.1 with Lemma 2.1, we prove Proposition 2.4.

PROOF OF PROPOSITION 2.4. To minimize notation, we set $\mathcal{G}_n = \mathcal{G}_n^{(a)}$, and let all constants depend implicitly on a :

(1) Using the partial order \preceq , we order the random variables $\{Z_{\mathbf{v}} : \mathbf{v} \in \mathcal{V}_n\}$ to have the $|\mathcal{V}_n| \times 1$ column vector \mathbf{Z}_n . Set $\mathbf{R}_{\Pi} = \sum_{\mathbf{v} \in \Pi} Y_{\mathbf{v}, \Pi} \mathbb{1}_{\{\mathbf{v}\}}$. Recalling the fact that $[\mathbf{A}(\mathcal{V}_n)]$ has zero diagonal entries,

$$(3.11) \quad \begin{aligned} \mathbb{E} \mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n &= 0, \\ \mathbb{E} \mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{R}_{\Pi} &= 0, \\ \mathbb{E} \mathbf{R}_{\Pi}^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{R}_{\Pi} &= 0, \end{aligned}$$

because each summand of all the above quadratic forms is product of independent random variables having mean zero. Observing that the summands $\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n$ are uncorrelated, and using (1) of Lemma 2.1,

$$\begin{aligned} \mathbb{E}[(\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n)^2] &= 4 \sum_{(i,u) \in \mathcal{V}_n} \sum_{(j,v) \in \mathcal{V}_n: (i,u) \preceq (j,v)} (j - i)^{-2} \\ &= 2 \|\mathbf{A}(\mathcal{V}_n)\|_F^2 \asymp n^2 \log n. \end{aligned}$$

Since W_n is the normalized version of $\mathbf{Z}_n^T [\mathbf{A}(\mathcal{V}_n)] \mathbf{Z}_n$, W_n has mean 0 and variance 1. Invoking Proposition 2.8 we get the desired bound.

(2) Using the partial order \preceq , define $|\mathcal{V}_n| \times 1$ column vector $\mathbf{X}_n = (X_{\mathbf{v}}, \mathbf{v} \in \mathcal{V}_n)$. Also define $\Upsilon_{\Pi} := \sum_{\mathbf{v} \in \Pi} \nu_{\mathbf{v}, \Pi} \mathbb{1}_{\{\mathbf{v}\}}$. Clearly,

$$\mathbf{X}_n \stackrel{d}{=} \begin{cases} \mathbf{Z}_n & \text{under } \mathbb{P}_0, \\ \mathbf{Z}_n + \mathbf{R}_{\Pi} + \Upsilon_{\Pi} & \text{under } \mathbb{P}_{1, \Pi}, \end{cases}$$

so

$$Q_n \stackrel{d}{=} \begin{cases} W_n & \text{under } \mathbb{P}_0, \\ W_n + U_n + v(\mathcal{G}_n) & \text{under } \mathbb{P}_{1,\Pi}, \end{cases}$$

where

$$U_n := 2\Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{Z}_n + 2\Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{R}_{\Pi} + 2\mathbf{R}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{Z}_n + \mathbf{n}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{R}_{\Pi}$$

and

$$h(\Pi, v_n) := \Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\Upsilon_{\Pi}.$$

It follows from (3.11) and the facts $\mathbb{E}\mathbf{Z}_n = \mathbb{E}\mathbf{R}_{\Pi} = \mathbf{0}$ that each summand of U_n has mean 0, so $\mathbb{E}U_n = 0$. Now we observe that:

- (a) $[\bar{\mathbf{A}}(\cdot)]$ has zero diagonal entries.
- (b) $\mathbb{E}[\mathbf{Z}_v W_n] = \mathbb{E}[Y_{v,\Pi} W_n] = 0$ for all $v \in \mathcal{Y}_n$, as components of \mathbf{Z}_n are independent with mean 0.
- (c) $\mathbb{E}[\mathbf{Z}_v \mathbf{Z}_{v'}^T Y_{u,\Pi} Y_{u',\Pi}] = \mathbb{E}[\mathbf{Z}_v \mathbf{Z}_{v'}^T \mathbf{Z}_u Y_{u',\Pi}] = 0$ for all $v, v', u, u' \in \mathcal{Y}_n$ satisfying $v \neq v', u \neq u'$, as $\{(Z_v, Y_{v,\Pi}) : v \in \mathcal{Y}_n\}$ are independent, and \mathbf{Z}_v and $Y_{v,\Pi}$ are uncorrelated for all $v \in \mathcal{Y}_n$.
- (d) $\mathbb{E}[\mathbf{Z}_n \mathbf{Z}_n^T] = I_{|\mathcal{Y}_n|}$ (the identity matrix), $\mathbb{E}[\mathbf{R}_{\Pi} \mathbf{R}_{\Pi}^T] \leq \text{Diag}(\mathbf{1}_{\Pi})$.

Using these observations, each of the summands of $W_n U_n$ has mean zero, so $\mathbb{E}W_n U_n = 0$. Also

$$\mathbb{E}[(\mathbf{R}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{Z}_n)^2] = \frac{1}{2} \|\mathbf{A}(\mathcal{Y}_n)\|_F^{-2} \sum_{v \in \Pi} \sum_{v' \in [\mathbf{v}]} \mathbb{E}Y_{v,\Pi}^2 \mathbb{E}Z_{v'}^2(A_{v,v'})^2,$$

$$\mathbb{E}[(\mathbf{R}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{R}_{\Pi})^2] = \frac{1}{2} \|\mathbf{A}(\mathcal{Y}_n)\|_F^{-2} \sum_{v \in \Pi} \sum_{v' \in [\mathbf{v}] \cap \Pi} \mathbb{E}Y_{v,\Pi}^2 \mathbb{E}Y_{v',\Pi}^2(A_{v,v'})^2.$$

Using Lemma 2.1(1), both terms in the above display are

$$\begin{aligned} &\leq \|\mathbf{A}(\mathcal{Y}_n)\|_F^{-2} \sum_{0 \leq i < j < n} \sum_{v:(i,u) \leftrightarrow (j,v) \text{ for some } u \in \Pi_i} (j-i)^{-2} \\ &\asymp \frac{1}{n^2 \log n} \sum_{0 \leq i < j < n} (j-i)^{-1} \asymp \frac{1}{n^2 \log n} \sum_{0 \leq i < n-1} \log(n-i) \asymp \frac{n \log n}{n^2 \log n} = 1/n. \end{aligned}$$

Combining these estimates with Lemma 2.1(4), (5) and using the Cauchy–Schwarz inequality,

$$\begin{aligned} \mathbb{E}U_n^2 &\leq 4[4\mathbb{E}(\Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{Z}_n)^2 + 4\mathbb{E}(\Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbf{R}_{\Pi})^2 + 5/n] \\ &= 16[\Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbb{E}(\mathbf{Z}_n \mathbf{Z}_n^T)[\bar{\mathbf{A}}(\mathcal{Y}_n)]\Upsilon_{\Pi} \\ &\quad + \Upsilon_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbb{E}(Y_{\Pi} Y_{\Pi}^T)[\bar{\mathbf{A}}(\mathcal{Y}_n)]\Upsilon_{\Pi}] + 20/n \end{aligned}$$

$$\begin{aligned} &\leq C\bar{v}^2[\mathbb{1}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]^2\mathbb{1}_{\Pi} + \mathbb{1}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\text{Diag}(\mathbb{1}_{\Pi})[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbb{1}_{\Pi}] + 1/n \\ &\leq C\bar{v}^2\left[\frac{1}{\log n} + \frac{\log n}{n}\right] + 1/n. \end{aligned}$$

(3) For each $i \in I$, define $\mathbf{v}_i := \arg \max_{\mathbf{v} \in \Pi} \nu_{\mathbf{v}, \Pi}$. If we let $\Pi(I) := \{v_i : i \in I\}$, then it is easy to see that $\nu(\mathcal{G}_n) \geq \bar{v}^2 \mathbb{1}_{\Pi(I)}^T[\bar{\mathbf{A}}(\mathcal{G}_n)]\mathbb{1}_{\Pi(I)}$. Since $|\Pi(I)| \geq n/2$, we can use Lemma 3.1 and conclude that there is a constant $c \in (0, 1)$ such that $\mathbb{1}_{\Pi(I)}^T[\bar{\mathbf{A}}(\mathcal{G}_n)]\mathbb{1}_{\Pi(I)} \geq cn \log n$. Combining this with Lemma 2.1(1), we get the desired lower bound.

For the upper bound, note that $\nu(\mathcal{G}_n) \leq \bar{v}^2 \mathbb{1}_{\Pi}^T[\bar{\mathbf{A}}(\mathcal{Y}_n)]\mathbb{1}_{\Pi}$, and use Lemma 2.1(1), (3). \square

Acknowledgment. The authors thank Hubert Lacoin for the reference to [2].

REFERENCES

- [1] ARIAS-CASTRO, E., CANDÈS, E. J., HELGASON, H. and ZEITOUNI, O. (2008). Searching for a trail of evidence in a maze. *Ann. Statist.* **36** 1726–1757. [MR2435454](#)
- [2] BERGER, Q. and LACOIN, H. (2017). The high-temperature behavior for the directed polymer in dimension 1 + 2. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 430–450. [MR3606747](#)
- [3] GAMKRELIDZE, N. G. and ROTAR', V. I. (1978). The rate of convergence in a limit theorem for quadratic forms. *Theory Probab. Appl.* **22** 394–397. [MR0443040](#)
- [4] GÖTZE, F. and TIKHOMIROV, A. N. (1999). Asymptotic distribution of quadratic forms. *Ann. Probab.* **27** 1072–1098. [MR1699003](#)
- [5] LACOIN, H. (2010). New bounds for the free energy of directed polymers in dimension 1 + 1 and 1 + 2. *Comm. Math. Phys.* **294** 471–503. [MR2579463](#)
- [6] ROTAR', V. I. (1974). Certain limit theorems for polynomials of degree two. *Theory Probab. Appl.* **18** 499–507. [MR0326803](#)
- [7] ROTAR', V. I. and SHERVASHIDZE, T. L. (1986). Some estimates for distributions of quadratic forms. *Theory Probab. Appl.* **30** 585–590. [MR0805309](#)
- [8] WHITTLE, P. (1960). Bounds for the moments of linear and quadratic forms in independent variables. *Theory Probab. Appl.* **5** 302–305. [MR0133849](#)

DEPARTMENT OF MATHEMATICS
CITY UNIVERSITY OF NEW YORK, CITY COLLEGE
160 CONVENT AVE
NEW YORK, NEW YORK 10031
USA
E-MAIL: shirshendu@ccny.cuny.edu
URL: <http://mathsci.ccnysites.cuny.edu/>

DEPARTMENT OF MATHEMATICS
WEIZMANN INSTITUTE OF SCIENCE
POB 26
REHOVOT 76100
ISRAEL
E-MAIL: ofer.zeitouni@weizmann.ac.il
URL: <http://www.wisdom.weizmann.ac.il/~zeitouni/>