

NUMBER OF ARITHMETIC PROGRESSIONS IN DENSE RANDOM SUBSETS OF $\mathbb{Z}/n\mathbb{Z}$

ROSS BERKOWITZ, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. We examine the behavior of the number of k -term arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$. We prove that if a set is chosen by including each element of $\mathbb{Z}/n\mathbb{Z}$ independently with constant probability p , then the resulting distribution of k -term arithmetic progressions in that set, while obeying a central limit theorem, does not obey a local central limit theorem. The methods involve decomposing the random variable into homogeneous degree d polynomials with respect to the Walsh/Fourier basis. Proving a suitable multivariate central limit theorem for each component of the expansion gives the desired result.

1. INTRODUCTION

Understanding the asymptotic behavior of sums of dependent random variables is a fundamental question in probability theory and combinatorics. One particular random variable that has received attention is the number of arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$. For any subset $S \subseteq \mathbb{Z}/n\mathbb{Z}$ we define $\mathbf{kAP}(S)$ to count the number of k -term arithmetic progressions contained entirely in the set S . The probability space is constructed by choosing a random set S by including each element of $\mathbb{Z}/n\mathbb{Z}$ independently at random with probability $p \in (0, 1)$, where p is a fixed constant not depending on n . The natural question therefore is to understand the distribution of $\mathbf{kAP}(S)$ as n grows.

It is not hard to show that $\mathbf{kAP}(S)$ obeys a central limit theorem. That is, if we set $\mu_n = \mathbb{E}[\mathbf{kAP}(S)]$ and $\sigma_n^2 = \text{Var}(\mathbf{kAP}(S))$, where S is chosen as before, then for any fixed a, b

$$\mathbb{P}\left[a \leq \frac{\mathbf{kAP}(S) - \mu_n}{\sigma_n} \leq b\right] = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{t^2}{2}\right) dx + o_{n,p}(1).$$

Given this Gaussian macroscopic behavior it is natural to guess that the distribution of \mathbf{kAP} is “smooth” and therefore nearby integers are approximately as likely as one another. In particular, one may conjecture that a local limit theorem estimating pointwise probabilities of $\mathbf{kAP}(S)$ that for any integer x

$$\mathbb{P}[\mathbf{kAP}(S) = x] \stackrel{?}{=} \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n)^2}{2\sigma_n^2}\right) + o\left(\frac{1}{\sigma_n}\right).$$

However the purpose of this note is to prove that this local limit theorem is in fact false and the distribution of $\mathbf{kAP}(S)$ oscillates wildly.

Theorem 1.1. *Fix $p \in (0, 1)$ and $k \geq 3$. Then for all sufficiently large n relatively prime to $(k-1)!$ there is a point x such that*

$$\left| \mathbb{P}[\mathbf{kAP}(S) = x] - \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n)^2}{2\sigma_n^2}\right) \right| = \Omega\left(\frac{1}{\sigma_n}\right),$$

where μ_n and σ_n in the statement are the expectation and standard deviation of $\mathbf{kAP}(S)$ and S is constructed by choosing each element of $\mathbb{Z}/n\mathbb{Z}$ independently at random with probability p .

Remark. This failure of the local central limit likely extends to $\gcd(n, (k-1)!) \neq 1$, however the proof details become more technical and therefore we restrict our attention to this case.

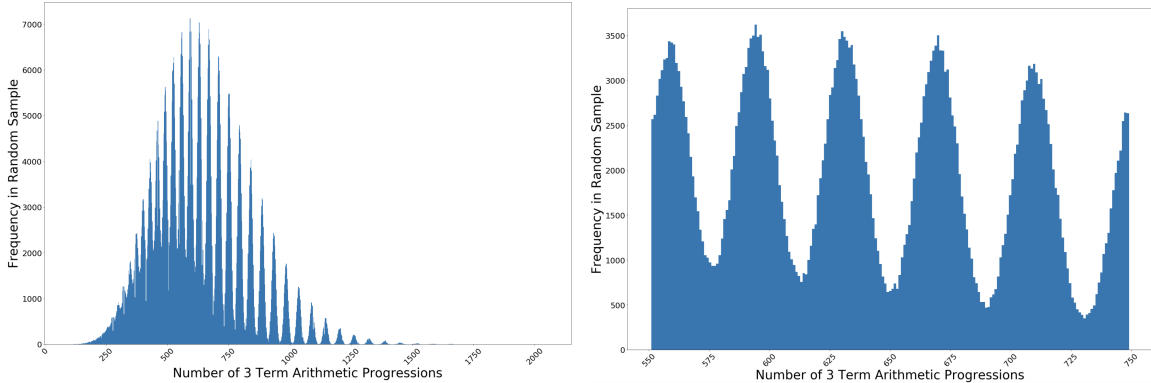


FIGURE 1. Histogram from sampling uniformly random subsets of $\mathbb{Z}/101\mathbb{Z}$ where 1,000,000 random samples were taken. While the the Gaussian-like distribution of $\mathbf{3AP}(S)$ is visible there are wild local fluctuations. The second picture on the right narrows the histogram to only looking at $550 \leq \mathbf{3AP}(S) \leq 750$ showing the local fluctuations in greater detail.

The first author discovered this failure of the local central limit theorem by sampling uniformly random subsets of $\mathbb{Z}/101\mathbb{Z}$ and counting the number of length 3 arithmetic progressions. This histogram of results may be found in Figure 1. Interestingly, it should be noted that subsequently and independently a study of Cai, Chen, Heller, and Tsegaye [CCHT] also conjectured that such a local limit theorem failed, but did not have a proof.

Related Work. Significant attention has been given to understanding the large deviation probability of $\mathbf{kAP}(S)$, particularly in the sparse set regime where $p \rightarrow 0$. For example, recently Warnke [War17], Bhattacharya, Ganguly, Shao, and Zhao [BGSZ20], and Harel, Mousset, and Samotij [HMS] found precise upper tail bounds for $\mathbf{kAP}(S)$ in the sparse regime, while Janson and Warnke [JW16] proved lower tail bounds. Additionally, Barhoumi-Andréani, Koch, and Liu [BAKL⁺19] proved a bivariate central limit theorem for $(\mathbf{mAP}(S), \mathbf{nAP}(S))$, understanding the joint distribution of the number of length m and n arithmetic progressions in sparse random sets.

Significant attention has also been focused on understanding local limit theorems in the analogous setting of $G(n, p)$. In particular work of Gilmer and Kopparty [GK16] proves a local central theorem for triangle counts, with Berkowitz giving improved bounds in the case of triangles [Berb] and then proving the analogous theorem for cliques in [Bera]. Furthermore for general connected subgraph counts in $G(n, p)$ almost optimal anti-concentration results are known due to the work of Fox, Kwan, and Sauermann [FKS]. Our work points to a certain degree of separation between a local central limit theorem and anti-concentration for polynomial functions of Bernoulli random variables as already suggested by Fox, Kwan, and Sauermann [FKS]. In particular, the random variable $\mathbf{kAP}(S)$ experimentally appears to satisfy anti-concentration (at the optimal scale with each point probability being at most $O(1/n^{3/2})$) but as we will prove it does not satisfy a local central limit theorem.

Outline of Paper. In Section 2 we compute the expansion of the \mathbf{kAP} in the p -biased Fourier basis. We then use this expansion to give a high level overview of our arguments. Sections 3-7 contain the main technical work of analyzing the asymptotic behavior of \mathbf{kAP} , and a more detailed overview of the argument can be found at the end of Section 3. The proof of the nonexistence of a local central limit theorem is in Section 8. Finally, we end with some outstanding questions left by our work in Section 9.

2. EXPANSION OF k -AP FUNCTION INTO p -BIASED BASIS AND OUTLINE OF THE ARGUMENT

We first expand the counting function of the number of k -APs into a p -biased Fourier basis. In order to do so define x_i to be indicator if the element i is in the subset of $\mathbb{Z}/n\mathbb{Z}$ which we are examining. Then we use the change of variables

$$y_i = \frac{x_i - p}{\sqrt{p(1-p)}}$$

and note that $\mathbb{E}[y_i] = 0$ and $\text{Var}[y_i] = 1$. Now the k -AP counting function is

$$\begin{aligned} \mathbf{kAP}(x) &= \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \prod_{i=0}^{k-1} x_{a+id} \\ &= \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \prod_{i=0}^{k-1} (y_{a+id} \sqrt{p(1-p)} + p) \\ &= \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{\ell=0}^k \sum_{S \in \binom{[k]}{\ell}} p^{k-|S|} \prod_{i \in S} (y_{a+id} \sqrt{p(1-p)}) \\ &= \sum_{\ell=0}^k \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{S \in \binom{[k]}{\ell}} p^{k-\frac{|S|}{2}} (1-p)^{\frac{|S|}{2}} \prod_{i \in S} y_{a+id}. \end{aligned}$$

Furthermore define

$$\mathbf{kAP}^\ell(y) = \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{S \in \binom{[k]}{\ell}} p^{k-\frac{\ell}{2}} (1-p)^{\frac{\ell}{2}} \prod_{i \in S} y_{a+id}.$$

The key idea is to note that \mathbf{kAP}^ℓ for $\ell = 1$ versus all higher values of ℓ live on different scales. Our main lemma will be to prove a quantitative convergence of these components to k appropriately scaled multivariate Gaussian and then using this analysis we will subsequently prove the desired failure of a local central limit theorem. For the sake of simplicity we also define

$$\overline{\mathbf{kAP}}^\ell(y) = \frac{1}{\sigma_\ell} \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{S \in \binom{[k]}{\ell}} \prod_{i \in S} y_{a+id}$$

where σ_ℓ is chosen so that $\text{Var}[\overline{\mathbf{kAP}}^\ell(z)] = 1$, if z is a vector of independent standard normals. In particular note σ_ℓ is independent of p . Note that these multivariate functions will be the central object of study and proving a sufficient strong result regarding their joint distribution will give the desired failure of a local central limit theorem. Finally define σ to be the variance of $\mathbf{kAP}(z)$.

Note that all the different functions defined here are multilinear, since $\text{gcd}(n, (k-1)!) = 1$.

Asymptotics of σ_ℓ . We note that $\sigma_\ell = \Theta_k(n)$ is easily computed for $\ell \neq 1$ and that $\sigma_1 = \Theta_k(n^{\frac{3}{2}})$. This follows immediately from the fact that any two elements of $\mathbb{Z}/n\mathbb{Z}$ lie in $O_k(1)$ k -term arithmetic progressions jointly.

Overview of the Main Arguments. Given the above expansion we are now in position to give a general overview of the proof. The argument is centered on demonstrating that the functions \mathbf{kAP}^1 and the remaining \mathbf{kAP}^ℓ fluctuate independently and on differing scales and then use this to deduce a failure of the local central limit theorem. The key claim is that $\{\mathbf{kAP}^\ell\}_{\ell=1, 3 \leq \ell \leq k}$, suitably normalized, approaches in distribution a set of independent Gaussian (along with quantitative bounds). In particular we prove the following result.

Theorem 2.1. For $g \in C^3(\mathbb{R}^{k-1})$, we have

$$|\mathbb{E}g(\overline{\mathbf{kAP}}^1(y), \overline{\mathbf{kAP}}^3(y), \dots, \overline{\mathbf{kAP}}^k(y)) - \mathbb{E}g(Z)| \lesssim_k \frac{M_2(g) + M_3(g)}{n^{1/2}},$$

where Z is a standard Gaussian random vector in \mathbb{R}^{k-1} .

In the notation above $M_2(g)$ and $M_3(g)$ are the maximal operator norms of the second and third order derivative tensors of g ; more informally these quantities simply measure the fluctuations in g . In order to prove the desired CLT result we first use a version of the Gaussian Invariance Principle which allows one to replace scaled Bernoulli's with Gaussians; this reduction appears in Section 3. This reduction, while not strictly necessary, simplifies the argument. Then we use the theory of exchangeable pairs to deduce the necessary central limit theorem, which constitutes Section 4. Roughly, the exchangeable pairs argument proceeds by using analyzing a single draw of Gaussian y 's for vector $(\overline{\mathbf{kAP}}^1(y), \overline{\mathbf{kAP}}^3(y), \dots, \overline{\mathbf{kAP}}^k(y))$ and analyzes what occurs if precisely one of the y at random is resampled. The method of exchangeable pairs allows one to deduce a quantitative central limit theorem from this perturbative analysis; however, a significant amount of effort is expended in verifying the necessary moment estimates. In particular, one consequence of the above analysis is that

$$\frac{\sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} z_a z_{a+d} z_{a+2d}}{\sqrt{\binom{n}{2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

if z_i are independent standard normals, since this corresponds to $\overline{\mathbf{kAP}}^3$ for $k = 3$. Specializing the analysis to this case may be useful for some readers. Note here that if z_i were not centered then the standard deviation jumps from $\Theta(n)$ to $\Theta(n^{3/2})$, and the corresponding central limit theorem is a easy consequence of the method of dependency graphs as demonstrated in [CCHT].

We next convert these results into a bound on Kolmogorov distance and then deduce the failure of a local central limit theorem using a sampling argument in Sections 7 and 8. Ultimately the failure of the local central theorem is derived essentially from the fact that \mathbf{kAP}^1 takes on values which are separated by $\Theta_{k,p}(n)$ from one another and the smearing which occurs due to the remaining components also lives on the scale of $\Theta_{k,p}(n)$. Thus it is not able to flatten this effect out. These two sections give one way of implementing this intuition.

3. REDUCTION TO GAUSSIAN ESTIMATE

We first use an invariance principle for multilinear polynomials. In order to do so we need to define the influence of a variable for a Boolean function and whether a random variable is hypercontractive. This step is not strictly speaking necessary, but does not weaken our bounds and allows us to establish the rest of the argument in a slightly cleaner form.

Definition 3.1. The influence of a variable x_i in a boolean function $F(x_1, \dots, x_n) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i$ is

$$\mathbf{Inf}_i[F] = \sum_{t \in S \subseteq [n]} a_S^2.$$

Definition 3.2. A random variable X is (p, q, ρ) -hypercontractive ($1 \leq p \leq q \leq \infty$ and $0 \leq \rho < 1$) if for all constants $a, b \in \mathbb{R}$ we have

$$\|a + \rho b X\|_q \leq \|a + b X\|_p.$$

Finally we need that the p -biased bit is hypercontractive with the appropriate constants. This follows from the following result in [O'D14].

Theorem 3.3. *If X is a mean zero, symmetric, discrete random variable with*

$$\lambda = \min_{x \in \text{Range}(X)} \mathbb{P}[X = x].$$

Then X is $(2, 3, \rho)$ -hypercontractive for $\rho = \frac{1}{\sqrt{q-1}} \cdot \lambda^{\frac{1}{2} - \frac{1}{q}}$.

We are now ready to define a version of the multifunction invariance principle; this version appears in [O'D14].

Theorem 3.4. *Let $F^{(1)}, \dots, F^{(d)}$ be formal n -variate multilinear polynomials each of degree at most $k \in \mathbb{N}$. Let x_1, \dots, x_n and y_1, \dots, y_n be independent \mathbb{R} -valued random variables such that $\mathbb{E}[x_t] = \mathbb{E}[y_t] = 0$ and $\mathbb{E}[x_t^2] = \mathbb{E}[y_t^2] = 1$. Assume each random variable x_t and y_t is $(2, 3, \rho)$ -hypercontractive. Then for any C^3 function $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfying $\|\partial^\beta \psi\|_\infty \leq C$ for all $|\beta| = 3$,*

$$\left| \mathbb{E}[\psi(F(x)) - \psi(F(y))] \right| \leq \frac{Cd^2}{3\rho^{3k}} \sum_{t=1}^n \sum_{j=1}^d \mathbf{Inf}_t[F^{(j)}]^{3/2}$$

For our application this will amount to the following theorem on the distribution of $\overline{\mathbf{kAP}}(y)$ against test functions.

Theorem 3.5. *Let y_i be defined as before and y'_i be standard normal random variables. Then for any C^3 function $\psi : \mathbb{R}^{k-1} \rightarrow \mathbb{R}$ satisfying $\|\partial^\beta \psi\|_\infty \leq C$ for all $|\beta| = 3$,*

$$\left| \mathbb{E}[\psi(\overline{\mathbf{kAP}}^1(y'), \overline{\mathbf{kAP}}^3(y'), \dots, \overline{\mathbf{kAP}}^k(y')) - \psi(\overline{\mathbf{kAP}}^1(y), \overline{\mathbf{kAP}}^3(y), \dots, \overline{\mathbf{kAP}}^k(y))] \right| \leq \frac{C_{k,p}}{n^{1/2}}$$

where the constant $C_{k,p}$ is linearly proportional to C .

Here we used that $\mathbf{Inf}_t[F^{(j)}] = O_{k,p}(1/n)$, which easily follows from the asymptotics of σ_ℓ along with the symmetry among the variables.

4. EXCHANGEABLE PAIRS

We now consider the joint distribution of $\{\overline{\mathbf{kAP}}^\ell(y_i)\}_{1,3 \leq \ell \leq k}$ where the y_i are now independent standard normals. We prove a quantitative result regarding its convergence to a multivariate normal using an application of exchangeable pairs. In order to state the version of exchange pairs, we will need we will first define a set of notations. Define for a real matrix

$$\langle A, B \rangle = \text{Tr}(A^T B)$$

and

$$\|A\|_{HS} = \sqrt{\text{Tr}(A^T A)} = \sqrt{\langle A, A \rangle}.$$

Furthermore define

$$\|A\|_{\text{op}} = \sup_{|v|=1, |w|=1} |\langle Av, w \rangle|$$

for matrices and similar for k -order forms

$$\|A\|_{\text{op}} = \sup_{|v_i|=1} |A(v_1, v_2, \dots, v_k)|.$$

Given this define the k^{th} derivative (tensor) operators for $f \in C^k(\mathbb{R}^n)$ as

$$\langle D^k f(x), (u_1, \dots, u_k) \rangle = \sum_{i_1, i_2, \dots, i_k \in [n]} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(u_1)_{i_1} \dots (u_k)_{i_k}$$

for vectors $u_1, \dots, u_k \in \mathbb{R}^n$. Finally define

$$M_r(g) = \sup_{x \in \mathbb{R}^n} \|D^r g(x)\|_{\text{op}}.$$

The last notion we need is that of exchangeable random variables.

Definition 4.1. X' and X are exchangeable random variables if (X', X) and (X, X') have the same distribution.

The key probability theoretic statement we will use is a multivariate version of exchangeable random variables for proving convergence to a Gaussian which in this form is due to Meckes [Mec09].

Theorem 4.2 ([Mec09]). *Let (X, X') be an exchangeable pair of random vectors in \mathbb{R}^d . Suppose that there is an invertible matrix Λ , and a random matrix E' such that*

- $\mathbb{E}[X' - X|X] = -\Lambda X$
- $\mathbb{E}[(X' - X)(X' - X)^T|X] = 2\Lambda + \mathbb{E}[E'|X]$.

Then for $g \in C^3(\mathbb{R}^d)$,

$$|\mathbb{E}g(X) - \mathbb{E}g(Z)| \leq \|\Lambda^{-1}\|_{op} \left[\frac{\sqrt{d}}{4} M_2(g) \mathbb{E}\|E'\|_{HS} + \frac{1}{9} M_3(g) \mathbb{E}|X' - X|^3 \right] \quad (1)$$

where Z is a standard Gaussian random vector in \mathbb{R}^d .

Note this is a simplification of the statement which appears in [Mec09] which is sufficient for our purposes. We now apply this to our setting where we set $y = (y_1, \dots, y_n)$ and $y' = (y_1, \dots, y'_I, \dots, y_n)$ where I is a uniformly random coordinate in $[n]$ and y_i, y'_I are independent standard normals. It obvious by definition that

$$W = \{\overline{\mathbf{kAP}}^\ell(y)\}_{\ell=1,3 \leq \ell \leq k}, W' = \{\overline{\mathbf{kAP}}^\ell(y')\}_{\ell=1,3 \leq \ell \leq k}$$

are exchangeable random variables. We stress that the $\ell = 2$ term is missing. Furthermore, our normalization of σ_ℓ has made it so that each coordinate of W has variance 1. Also, $\mathbb{E}[W] = 0$ since $\overline{\mathbf{kAP}}^\ell(y)$ is multi-linear as we have $\gcd(n, (k-1)!) = 1$. We first compute the matrix Λ in the case of (W, W') .

Proposition 4.3. *Let W, W' be defined as above. Then*

$$\mathbb{E}[W' - W|W] = -\mathbf{diag}\left(\frac{i}{n}\right)_{i=1,3 \leq i \leq k} W.$$

Proof. Note that the ℓ^{th} coordinate of $\mathbb{E}[W' - W|y]$ is

$$\begin{aligned} & \frac{1}{n} \sum_{m=1}^n \mathbb{E}[\overline{\mathbf{kAP}}^\ell(y_1, \dots, y'_m, y_{m+1}, \dots) - \overline{\mathbf{kAP}}^\ell(y)|y] \\ &= \frac{1}{n} \sum_{m=1}^n \mathbb{E}[\overline{\mathbf{kAP}}^\ell(y_1, \dots, 0, y_{m+1}, \dots) - \overline{\mathbf{kAP}}^\ell(y)|y] \\ &= -\frac{1}{n} \mathbb{E}\left[\sum_{m=1}^n \frac{1}{\sigma_\ell} \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{S \in \binom{[k]}{\ell}} \prod_{i \in S, m \in a+dS} y_{a+id} | y \right] \\ &= -\frac{1}{n} \left(\mathbb{E}[\ell(\overline{\mathbf{kAP}}^\ell(y))|y] \right) = -\frac{\ell}{n} \left(\overline{\mathbf{kAP}}^\ell(y) \right) \end{aligned}$$

and the proposition follows upon taking a conditional expectation with respect to W . The first equality follows because, conditional on the index $I = m$ which was removed, y'_m is independent from everything else and has mean zero and $\overline{\mathbf{kAP}}^\ell$ is multilinear. \square

Now to apply Theorem 4.2 we will simply take

$$E' = \mathbb{E}[(W' - W)(W' - W)^T - 2\Lambda|y],$$

which clearly satisfies the necessary hypothesis. To apply Theorem 4.2 to W, W' , we now see it suffices to bound $\mathbb{E}[\|E'\|_{HS}]$ and $\mathbb{E}|W' - W|^3$. Noting that $\|\Lambda^{-1}\|_{\text{op}} = n$, we will want bounds of the form $O_k(n^{-\frac{3}{2}})$ for each of these two quantities.

It is worth studying the diagonal terms more carefully. We have

$$\mathbb{E}[(W'_\ell - W_\ell)^2] = \mathbb{E}[2W_\ell^2 - 2W_\ell W'_\ell] = \mathbb{E}[2W_\ell \mathbb{E}[W_\ell - W'_\ell | W_\ell]] = \frac{2\ell}{n} \mathbb{E}[W_\ell^2] = \frac{2\ell}{n}$$

by exchangeability, conditional expectations, Proposition 4.3, and the normalization of W_ℓ . Therefore

$$(E')_{\ell, \ell} = \mathbb{E}[(W'_\ell - W_\ell)^2 | y] - \mathbb{E}[(W'_\ell - W_\ell)^2]$$

for $1 \leq \ell \leq k$ and $\ell \neq 2$.

Computing E' . We begin by simply computing E' entry by entry. For this we define the further refinement

$$\overline{\mathbf{kAP}}^{\ell, t}(y) = \frac{1}{y_t \sigma_\ell} \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \sum_{S \in \binom{[k]}{\ell}} \prod_{i \in S, t \in a+dS} y_{a+id}.$$

Note that the above in theory is not defined when y_t is 0, but really we are simply removing the term y_t from all products in the summation so this can be extended in the obvious way. Less formally this is the sum $\overline{\mathbf{kAP}}^\ell$ with all the terms containing the term y_t with y_t factored out. Using this notation it follows easily that the nondiagonal entries are

$$(E')_{i, j} = \frac{1}{n} \sum_{s \in [n]} \overline{\mathbf{kAP}}^{i, s}(y) \overline{\mathbf{kAP}}^{j, s}(y) (y_s^2 + 1)$$

and the diagonal entries are

$$(E')_{i, i} = -\frac{2i}{n} + \frac{1}{n} \sum_{s \in [n]} (y_s^2 + 1) \overline{\mathbf{kAP}}^{i, s}(y)^2.$$

Here we used that each random variable y_i has mean zero and variance one, and same for its replacement y'_i . In fact, the normalization of σ_ℓ implies that $\mathbb{E}[(E')_{i, i}] = 0$, although this is not obvious by direct computation. We can see this from cross-comparison with the earlier expression for $(E')_{\ell, \ell}$.

We now proceed further into the computational abyss and consider

$$\text{Tr}(E' E'^T) = \sum_{i, j} (E')_{i, j}^2.$$

We will need a bound on

$$\mathbb{E} \left[\sqrt{\text{Tr}(E' E'^T)} \right] \leq \sqrt{\mathbb{E}[\text{Tr}(E' E'^T)]}$$

which has decay properties of the form $O_k(n^{-3/2})$. Thus it suffices to prove a bound of the form $O_k(n^{-3})$ for

$$\mathbb{E} \left[\sum_{i, j} (E')_{i, j}^2 \right].$$

Note that it suffices to prove such a bound for each individual summands as there are k^2 such summands and the result will follow.

5. BOUNDING E'

5.1. **Non-diagonal terms.** We will first consider the case where $i \neq j$. Without loss of generality let $i < j$. Thus, since $i \geq 1$ and $j \neq 2$, we have $j \geq 3$.

We first define an even further refinement of our polynomials,

$$\overline{\mathbf{kAP}}^{\ell,t,S}(y) = \frac{1}{y_t \sigma_\ell} \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} \prod_{i \in S, t \in a+dS} y_{a+id}.$$

Note that

$$(E')_{i,j} = \frac{1}{n} \sum_{S \in \binom{[k]}{i}} \sum_{T \in \binom{[k]}{j}} \sum_{v \in [n]} \overline{\mathbf{kAP}}^{i,v,S}(y) \overline{\mathbf{kAP}}^{j,v,T}(y) (y_v^2 + 1)$$

and thus it suffices to prove that for all $S \in \binom{[k]}{i}$ and $T \in \binom{[k]}{j}$ we have

$$\mathbb{E} \left[\left(\frac{1}{n} \sum_{v \in [n]} \overline{\mathbf{kAP}}^{i,v,S}(y) \overline{\mathbf{kAP}}^{j,v,T}(y) (y_v^2 + 1) \right)^2 \right] = O_k(n^{-3})$$

by Cauchy-Schwarz. Using Cauchy-Schwarz again, it suffices to instead prove that

$$\mathbb{E} \left[\left(\sum_{v \in [n]} \overline{\mathbf{kAP}}^{i,v,S}(y) \overline{\mathbf{kAP}}^{j,v,T}(y) \right)^2 + \left(\sum_{v \in [n]} \overline{\mathbf{kAP}}^{i,v,S}(y) \overline{\mathbf{kAP}}^{j,v,T}(y) y_v^2 \right)^2 \right] = O_k(n^{-1}).$$

The key idea is that when expanded as a polynomial in the y 's, the inside of the expectation will have most terms contain an odd power of some y_i , which leads to a zero contribution as y_i is a standard normal. Every nonzero term contributes an amount bounded by $O_k(1) \cdot (\sigma_i \sigma_j)^{-2}$ as the exponents are bounded. Note that the parities of the exponents are unchanged between the first and second terms, so we use the second out of convenience.

A term of the second, when expanded out, amounts to choosing $v_1, v_2 \in [n]$ and then $a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{Z}/n\mathbb{Z}$ and $d_{11}, d_{12}, d_{21}, d_{22} \in [n/2]$ such that $v_c \in a_{c1} + d_{c1}S$ and $v_c \in a_{c2} + d_{c2}T$ for $c \in \{1, 2\}$. Let A_{ct} for $k, t \in \{1, 2\}$ be the sets thus formed (there are no self-intersections since $\gcd(n, (k-1)!) = 1$). Let the multiset A equal the union, with repetition, of all A_{ct} . Note $|A_{c1}| = i$ and $|A_{c2}| = j$, so that $|A| = 2(i+j)$.

Claim 5.1. *Given an initial choice of three of the (a_{ct}, d_{ct}) , there are $O_k(1)$ ways to choose the remaining pair so as to be compliant with the condition that every element in A appears with even parity.*

Proof. After canceling we can see what parities the remaining set must have at each value of $\mathbb{Z}/n\mathbb{Z}$, which precisely determines it. (Recall that each is a set rather than a multiset because $\gcd(n, (k-1)!) = 1$.) Then there are $O_k(1)$ ways to choose the (a_{ct}, d_{ct}) given what the set must be. \square

First consider the case where A_{11}, A_{12} do not intersect at a place other than v_1 . This means they together hit $i+j-2$ distinct values once, and v_1 twice. We see this immediately implies that A_{21} and A_{22} , after removing v_2 from each, must hit precisely these $i+j-2$ distinct values. Since $j \geq 3$, this means that the value of v_2 is determined up to $O_k(1)$ choices by looking at the possible ways A_{22} hits these $i+j-2$ values in $j-1 \geq 2$ places. Therefore we see there are $O_k(1)$ choices of $a_{21}, d_{21}, a_{22}, d_{22}$ after selecting v_1, d_{11}, d_{12} , which means we have $O(n)^3 \cdot O_k(1)$ total choices (as there are $O_k(1)$ choices of a_{11}, a_{12} given that information).

The analysis is similar if A_{21}, A_{22} do not intersect other than at v_2 . So now we consider the case where A_{11}, A_{12} intersect at a place other than v_1 and same for A_{21}, A_{22} . After choosing v_1, v_2 with $O(n^2)$ choices, we claim there are $O_k(1)$ ways to finish choosing. First note that A_{11}, A_{12} have $O_k(n)$ possibilities together, and after that there are $O_k(1)$ choices for whichever of A_{21}, A_{22} intersects $A_{11} \cup A_{12}$ (which must happen as these sets A_{11}, A_{12} cannot cancel each other out, being

of differing size). And by Claim 5.1 there are $O_k(1)$ ways to choose the last one. This gives $O_k(n^3)$ once more.

In total, we have $O_k(n^3)$ terms that are not zero in the expectation. This yields a total contribution of $O_k(n^3) \cdot (\sigma_i \sigma_j)^{-2} = O_k(n^{-1})$, as desired.

5.2. Diagonal terms. Now we consider the diagonal terms $(E_{\ell,\ell})'$. From Section 4, we have

$$\mathbb{E}[(E'_{\ell,\ell})^2] = \mathbb{E}[(\mathbb{E}[(W'_\ell - W_\ell)^2|y] - \mathbb{E}[(W'_\ell - W_\ell)^2])^2] = \text{Var}(\mathbb{E}[(W'_\ell - W_\ell)^2|y]),$$

where the variance in the final line is over the randomness of the standard normals y . Now

$$\mathbb{E}[(W'_\ell - W_\ell)^2|y] = \mathbb{E}_{I,y'}[(y'_I - y_I)^2 \overline{\mathbf{kAP}}^{\ell,I}(y)^2] = \frac{1}{n} \sum_{i=1}^n (1 + y_i^2) \overline{\mathbf{kAP}}^{\ell,i}(y)^2.$$

If $\ell = 1$, $\overline{\mathbf{kAP}}^{\ell,i}(y)$ is a constant of size $\Theta_k(n^{-\frac{1}{2}})$. We end up with a bound of quality $O_k(n^{-3})$ trivially. Now let $\ell \geq 3$, recalling $\ell \neq 2$. We have

$$\mathbb{E}[(E'_{\ell,\ell})^2] = \frac{1}{n^2 \sigma_\ell^4} \sum_{1 \leq i, j \leq n} \text{Cov} \left[(1 + y_i^2) (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,i}(y))^2, (1 + y_j^2) (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,j}(y))^2 \right].$$

We first deal with the $i = j$ terms. They are bounded by $\mathbb{E}[(1 + y_i^2)^2 (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,i}(y))^4]$. Note that this value is independent of i , so we let $i = 0$ (which is the same as $i = n$). We adopt a similar method as before. It suffices to show that there are $O_k(n^2)$ terms of this $i = 0$ value that, when expanded, yield a nonzero expectation value. This is since there are n such terms and as $\sigma_\ell = \Theta_k(n)$, which would lead to an overall contribution of the desired size $O_k(n^{-3})$ to $\mathbb{E}[(E'_{\ell,\ell})^2]$.

We use the $\overline{\mathbf{kAP}}^{\ell,i,S}$ refinement from before. By Hölder's inequality, it suffices to bound each $\mathbb{E}[(1 + y_0^2)^2 (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,0,S}(y))^4]$. We also can choose an exponent of y_0 from the initial term, but it does not affect parities of exponents and is of constant order so we ignore this and assume we have the y_0^4 term for simplicity.

Each term within this sum is chosen via $d_1, \dots, d_4 \in [n/2]$ and offsets (e.g. which term equals the y_0 term that is being divided in $\overline{\mathbf{kAP}}^{\ell,0,S}$), each of which contributes terms of the form $y_{d_j} S_j$, where S_j is one of $O_k(1)$ many shifts of S that contains 0. Now for any valid tuple (d_1, \dots, d_6) , make a graph on vertex set $\{1, \dots, 6\}$, with i, j connected if $d_i S_i$ and $d_j S_j$ intersect other than at 0.

Given such a graph, we claim that there are $O_k(n)$ ways to choose the ds associated to a connected component of this graph in a manner compatible with the graph. Indeed, we find that, for example, if d_1, d_2, d_3 are connected, then choosing d_1 will fix the value say of cd_2 for some $c \in [k]$, which yields $O_k(1)$ possible values of d_2 , and then $O_k(1)$ possible values of d_3 similarly.

Furthermore, in a graph associated to a valid tuple (d_1, \dots, d_6) , i.e., one with even powers of the ys , we must have no disconnected vertices. Indeed, since the vertex is disconnected, the associated value of d_i must give rise to a multiset $d_i S \pmod{n}$ which must cancel out all of its contributions to the ys . This is impossible as $\gcd(n, (k-1)!) = 1$.

Finally, we have at most $\frac{4}{2} = 2$ connected components of non-isolated vertices, each of which have $O_k(n)$ ways to choose the ds . This yields an upper bound of $O_k(n^2)$, as desired.

Now we consider $i \neq j$. Again, by translation invariance we see the value only depends on $j - i$. Therefore it suffices to show for $i \neq 0$ that

$$\text{Cov} \left[(1 + y_0^2) (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,0}(y))^2, (1 + y_i^2) (\sigma_\ell \overline{\mathbf{kAP}}^{\ell,i}(y))^2 \right] = O_k(n)$$

since there are around n^2 total terms. We will show that

$$\text{Cov} \left[(\sigma_\ell y_0 \overline{\mathbf{kAP}}^{\ell,0}(y))^2, (\sigma_\ell y_i \overline{\mathbf{kAP}}^{\ell,i}(y))^2 \right] = O_k(n).$$

Again, the other four cases that need to be verified will be essentially identical, since we are just removing even exponent terms. Consider

$$\mathbb{E}[(\sigma_{\ell} y_0 \overline{\mathbf{kAP}}^{\ell,0}(y))^2 (\sigma_{\ell} y_i \overline{\mathbf{kAP}}^{\ell,i}(y))^2] = \sum_{S,T \in \binom{[k]}{\ell}} \sum_{a,b \in (\mathbb{Z}/n\mathbb{Z})^2} \sum_{\substack{d,e \in [n/2]^2 \\ 0 \in a_j + d_j S_j \\ i \in b_j + e_j T_j}} \mathbb{E} \left[\prod_{j=1}^2 \prod_{s_j \in S_j} y_{a_j + d_j s_j} \prod_{t_j \in T_j} y_{b_j + e_j t_j} \right].$$

We claim that the amount of nonzero terms other than those with $a_1 + d_1 S_1 = a_2 + d_2 S_2 = A$ and $b_1 + e_1 T_1 = b_2 + e_2 T_2 = B$ and $A \cap B = \emptyset$ is $O_k(n)$. Indeed, this is a similar argument as in the non-diagonal term case. The analogue of Claim 5.1 immediately follows, and similar arguments to earlier deal with the case that $a_1 + d_1 S_1, a_2 + d_2 S_2$ intersect only at 0. Indeed, if so, then they must hit $2\ell - 2$ distinct values once and 0 twice. Then $b_1 + e_1 T_1, b_2 + e_2 T_2$, after removing j from each, must precisely hit those $2\ell - 2$ values, each hitting $\ell - 1$ of them. As $\ell \geq 3$, we see that this pins down what b_1, e_1, b_2, e_2 are up to $O_k(1)$ possibilities, as linear combinations of the d_i (say after fixing which positions of $a_i + d_i S_i$ equal 0). Furthermore, we can compute j as a linear combination of the d_i . Since $j \neq 0$, it must be a nonzero linear combination. This pins down (d_1, d_2) to $O_k(n)$ possible values. Thus we have $O_k(n) \cdot O_k(1) = O_k(n)$ total possibilities. Furthermore, we have a similar analysis for the case where $b_1 + e_1 T_1, b_2 + e_2 T_2$ intersect only at j .

Now we consider the case where $a_1 + d_1 S_1, a_2 + d_2 S_2$ intersect other than at 0, and similar for $b_i + e_i T_i$. Then there are $O_k(n)$ choices for (a_1, d_1) , and then $O_k(1)$ choices for (a_2, d_2) , and if we are assuming $a_1 + d_1 S_1 \neq a_2 + d_2 S_2$, the terms $b_i + e_i T_i$ must hit one of these. There are thus $O_k(1)$ possibilities for the value of the set that does hit these values, and then $O_k(1)$ possibilities for the other set. Thus we have $O_k(n)$ choices once more, under the hypothesis that $a_1 + d_1 S_1 \neq a_2 + d_2 S_2$. Similar analysis holds if instead we assume $b_1 + e_1 T_1 \neq b_2 + e_2 T_2$. In the case where neither holds, so that $a_1 + d_1 S_1 = a_2 + d_2 S_2 = A$ and $b_1 + e_1 T_1 = b_2 + e_2 T_2 = B$, if we assume that $A \cap B \neq \emptyset$ then again we have a bound of $O_k(n)$,

The only remaining terms are those that have $a_1 + d_1 S_1 = a_2 + d_2 S_2 = A$ and $b_1 + e_1 T_1 = b_2 + e_2 T_2 = B$ and $A \cap B = \emptyset$, as desired.

Now we consider

$$\mathbb{E}[(\sigma_{\ell} y_0 \overline{\mathbf{kAP}}^{\ell,0}(y))^2] = \mathbb{E}[(\sigma_{\ell} y_i \overline{\mathbf{kAP}}^{\ell,i}(y))^2] = \sum_{S \in \binom{[k]}{\ell}} \sum_{a \in (\mathbb{Z}/n\mathbb{Z})^2} \sum_{\substack{d \in [n/2]^2 \\ 0 \in a_j + d_j S_j}} \mathbb{E} \left[\prod_{j=1}^2 \prod_{s_j \in S_j} y_{a_j + d_j s_j} \right].$$

We claim that the only nonzero terms are those with $a_1 + d_1 S_1 = a_2 + d_2 S_2$. Indeed, they are both sets as $\gcd(n, (k-1)!) = 1$, and if they are not equal then there is some element with an odd exponent.

Now, putting it all together, we see that

$$\text{Cov} \left[(1 + y_0^2) (\sigma_{\ell} \overline{\mathbf{kAP}}^{\ell,0}(y))^2, (1 + y_i^2) (\sigma_{\ell} \overline{\mathbf{kAP}}^{\ell,i}(y))^2 \right]$$

has a contribution of $O_k(n)$ terms in its $\mathbb{E}[XY]$ portion which we bound by $O_k(n)$. Otherwise it only has terms corresponding to $a_1 + d_1 S_1 = a_2 + d_2 S_2 = A$ and $b_1 + e_1 T_1 = b_2 + e_2 T_2 = B$ and $A \cap B = \emptyset$. Thus the expectations factor into a product of $\mathbb{E}[\prod_{a \in A} y_a^2]$ and $\mathbb{E}[\prod_{b \in B} y_b^2]$. This is canceled by the terms described by

$$\mathbb{E}[(\sigma_{\ell} y_0 \overline{\mathbf{kAP}}^{\ell,0}(y))^2] \mathbb{E}[(\sigma_{\ell} y_i \overline{\mathbf{kAP}}^{\ell,i}(y))^2].$$

There is one catch: the terms in this latter product of expectations $\mathbb{E}[\prod_{a \in A} y_a^2] \mathbb{E}[\prod_{b \in B} y_b^2]$ which have $A \cap B \neq \emptyset$ are not canceled in the $\mathbb{E}[XY]$ term. However, we see that there are $O_k(n) \cdot O_k(1)$ of them, for after choosing whichever of $O_k(n)$ values for A that we want, we have $O_k(1)$ choices

of B that both go through j and intersect A (unless A goes through j , but then it only has $O_k(1)$ choices and B has $O_k(n)$ choices in this case).

So, overall, the covariance is indeed $O_k(n)$, as desired.

6. BOUNDING MOMENTS

For the second part we need to prove that $\mathbb{E}[|W - W'|^3]$ is of size $O_k(n^{-3/2})$ which is the same as

$$\mathbb{E}\left[\sum_{\ell=1,3\leq\ell\leq k} |W^{(\ell)} - W'^{(\ell)}|^2\right]^{\frac{3}{2}} = O_k(n^{-\frac{3}{2}}).$$

Therefore it in fact suffices to prove that

$$\mathbb{E}[|W^{(\ell)} - W'^{(\ell)}|^3] = O_k(n^{-\frac{3}{2}}),$$

using Hölder's inequality. For $\ell = 1$ this is trivial, using $\sigma_1 = \Theta_k(n^{3/2})$. Now let $\ell > 1$ and note that

$$\begin{aligned} \mathbb{E}[(|W^{(\ell)} - W'^{(\ell)}|)^3] &= \mathbb{E}[|\overline{\mathbf{kAP}}^\ell(y) - \overline{\mathbf{kAP}}^\ell(y')|^3] \\ &= \mathbb{E}_{t \in \mathbb{Z}/n\mathbb{Z}} \mathbb{E}[|y_t - y'_t|^3 |\overline{\mathbf{kAP}}^{\ell,t}(y)|^3] \\ &= \mathbb{E}[|y_0 - y'_0|^3 |\overline{\mathbf{kAP}}^{\ell,0}(y)|^3] \\ &\leq \mathbb{E}[|y_0 - y'_0|^6]^{\frac{1}{2}} \mathbb{E}[|\overline{\mathbf{kAP}}^{\ell,0}(y)|^6]^{\frac{1}{2}} \\ &= \sqrt{120} \mathbb{E}[\overline{\mathbf{kAP}}^{\ell,0}(y)^6]^{\frac{1}{2}}. \end{aligned}$$

We now use the $\overline{\mathbf{kAP}}^{\ell,t,S}(y)$ refinement as in Section 5.1. Using Hölder's inequality again, it suffices to prove that

$$\mathbb{E}[\overline{\mathbf{kAP}}^{\ell,0,S}(y)^6] = O_k(n^{-3}).$$

We adopt a similar method to Section 5. Since $\sigma_\ell = \Theta_k(n)$, it amounts to showing there are $O_k(n^3)$ terms of $\mathbb{E}[\overline{\mathbf{kAP}}^{\ell,0,S}(y)^6]$ which have a nonzero contribution, i.e., even exponents of the y 's.

Each term within this sum is chosen via $d_1, \dots, d_6 \in [n/2]$, each of which contributes the terms $y_{d_j S}$. Now for any valid tuple (d_1, \dots, d_6) , make a graph on vertex set $\{1, \dots, 6\}$, with i, j connected if $d_i S$ and $d_j S$ intersect.

Each term within this sum is chosen via $d_1, \dots, d_6 \in [n/2]$ and offsets (e.g. which term equals the y_0 term that is being divided in $\overline{\mathbf{kAP}}^{\ell,0,S}$), each of which contributes terms of the form $y_{d_j T_j}$, where S_j is a set which is one of $O_k(1)$ many shifts of S that contain 0, and $T_j = S_j - \{0\}$. Now for any valid tuple (d_1, \dots, d_6) , make a graph on vertex set $\{1, \dots, 6\}$, with i, j connected if $d_i T_i$ and $d_j T_j$ intersect.

Given such a graph, we claim that there are $O_k(n)$ ways to choose the d s associated to a connected component of this graph in a manner compatible with the graph. Indeed, we find that, for example, if d_1, d_2, d_3 are connected, then choosing d_1 will fix the value say of cd_2 for some $c \in [k]$, which yields $O_k(1)$ possible values of d_2 , and then $O_k(1)$ possible values of d_3 similarly.

Furthermore, in a graph associated to a valid tuple (d_1, \dots, d_6) , i.e., one with even powers of the y s, we must have no disconnected vertices. Indeed, since the vertex is disconnected, the associated value of d_i must give rise to a multiset $d_i S \pmod{n}$ which must cancel out all of its contributions to the y 's. This is impossible as $\gcd(n, (k-1)!) = 1$.

Finally, we have at most $6/2 = 3$ connected components of non-isolated vertices, each of which have $O_k(n)$ ways to choose the d 's. This yields an upper bound of $O_k(n^3)$, as desired.

Conclusion. Putting the various estimates together we have proved the following result, which is a restatement of Theorem 2.1.

Theorem 6.1. *For $g \in C^3(\mathbb{R}^{k-1})$, we have*

$$|\mathbb{E}g(W) - \mathbb{E}g(Z)| \lesssim_k \frac{M_2(g) + M_3(g)}{n^{1/2}},$$

where Z is a standard Gaussian random vector in \mathbb{R}^{k-1} .

7. CONVERSION TO BOUND IN KOLMOGOROV DISTANCE

We now convert this test function bound into a bound on the cumulative distribution function.

Lemma 7.1. *For any a, b we have that*

$$\mathbb{P}\left[\overline{\mathbf{kAP}}^1(y) < a, \frac{\sum_{i=3}^k \sigma_i p^{k-\frac{i}{2}} (1-p)^{\frac{i}{2}} \overline{\mathbf{kAP}}^i(y)}{\sqrt{\sum_{i=3}^k \sigma_i^2 p^{2k-i} (1-p)^i}} < b\right] = \frac{1}{2\pi} \int_{-\infty}^a \int_{-\infty}^b e^{-\frac{(x^2+y^2)}{2}} dx dy + O_{k,p}(n^{-1/8}).$$

Proof. The key idea is to take ϕ_ℓ which is a smooth function which is 1 on $(-\infty, \ell]$, 0 on $[\ell + \epsilon, \infty)$, has second derivative bounded by $O(\epsilon^{-2})$, and third derivative bounded by $O(\epsilon^{-3})$. Furthermore let ϕ_ℓ be in $[0, 1]$ over the entire domain. Given this define $\gamma_{a,b}(x, y) = \phi_a(x)\phi_b(y)$. It follows from $\sigma_i = \Theta_k(n)$ for all $3 \leq i \leq k$ that

$$\Psi_{a,b}(y_1, y_3, \dots, y_k) = \gamma_{a,b}\left(y_1, \frac{\sum_{i=3}^k \sigma_i p^{k-\frac{i}{2}} (1-p)^{\frac{i}{2}} y_i}{\sqrt{\sum_{i=3}^k \sigma_i^2 p^{2k-i} (1-p)^i}}\right)$$

has $M_2(\Psi_{a,b}) = \Theta_{k,p}(\epsilon^{-2})$, $M_3(\Psi_{a,b}) = \Theta_{k,p}(\epsilon^{-3})$, and $\|\partial^\beta \Psi_{a,b}\| = \Theta_{k,p}(\epsilon^{-3})$ for all $|\beta| = 3$. The key point of course is that for standard Gaussians z_i we have

$$\mathbb{E}[\Psi_{a,b}(z_1, z_3, \dots, z_k)] = \frac{1}{2\pi} \int_{-\infty}^a \int_{-\infty}^b e^{-\frac{(x^2+y^2)}{2}} dx dy + O_k(\epsilon)$$

and by Theorems 2.1 and 3.5 it follows that

$$\mathbb{E}[\Psi_{a,b}(\overline{\mathbf{kAP}}^1(y), \overline{\mathbf{kAP}}^3(y), \dots, \overline{\mathbf{kAP}}^k(y))] = \mathbb{E}[\Psi_{a,b}(z_1, z_3, \dots, z_k)] + O_{k,p}\left(\frac{\epsilon^{-2} + \epsilon^{-3}}{n^{1/2}}\right).$$

Choosing $\epsilon = n^{-1/8}$, we obtain the desired right side as an upper bound, noting that $\Psi_{a,b}$ dominates the desired indicator function. A lower bound is obtained in an analogous manner and the result then follows. \square

8. PROOF OF THE FAILURE OF LOCAL CENTRAL LIMIT THEOREM

In this section we prove that \mathbf{kAP} does not obey a local central limit theorem. Specifically:

Theorem 1.1. *Fix $p \in (0, 1)$ and $k \geq 3$. Then for all sufficiently large n relatively prime to $(k-1)!$ there is a point x such that*

$$\left| \mathbb{P}[\mathbf{kAP}(z) = x] - \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n)^2}{2\sigma_n^2}\right) \right| = \Omega\left(\frac{1}{\sigma_n}\right),$$

where μ_n and σ_n in the statement are the expectation and standard deviation of $\mathbf{kAP}(z)$ and z is sampled i.i.d. with probability p .

We first give a high level overview of the proof. The proof proceeds by building two sets L_α and L_β of equal size both close enough to the mean of \mathbf{kAP} so that were there to be a local limit theorem for \mathbf{kAP} we would necessarily have

$$\mathbb{P}[\mathbf{kAP} \in L_\alpha] \approx \mathbb{P}[\mathbf{kAP} \in L_\beta] \approx \frac{|L_\alpha|}{\sigma_n \sqrt{2\pi}}.$$

However, as a result of Lemma 7.1 we will be able to compute $\mathbb{P}[\mathbf{kAP} \in L_\alpha] \propto f_\delta(\alpha)$, and $\mathbb{P}[\mathbf{kAP} \in L_\beta] \propto f_\delta(\beta)$, where f_δ is as defined in Lemma 8.1 below.

The rough idea in building L_α and L_β is to note that $\mathbf{kAP}^1 + \mathbf{kAP}^2$ takes values very near to a lattice $G\mathbb{Z}$ where $G = \Theta(n)$. Meanwhile $\mathbf{kAP}^{>2}$ has standard deviation $\Theta(n)$. L_α (and L_β) will roughly correspond to the event that $\mathbf{kAP} \approx \alpha G \pmod{G}$ ($\approx \beta G \pmod{G}$ respectively), and we can use our joint central limit theorem for \mathbf{kAP}^1 and $\mathbf{kAP}^{>2}$ to show that $\mathbb{P}[\mathbf{kAP} \approx \alpha G \pmod{G}] \propto f_\delta(\alpha)$ for some choice of δ . But since $f_\delta(\alpha) \neq f_\delta(\beta)$ we will conclude that $\mathbb{P}(L_\alpha) \neq \mathbb{P}(L_\beta)$ disproving any chance that \mathbf{kAP} obeys a local central limit theorem.

Before beginning our proof, we define f_δ and prove it is nonconstant.

Lemma 8.1. *The function*

$$f_\delta(x) = \sum_{\lambda \in \mathbb{Z}} e^{-\frac{(x-\lambda)^2}{\delta}}$$

is not the constant function for any value $\delta > 0$. Furthermore for any $C > 0$ there is a constant $D > 0$ so that if $\delta < C$ then for some pair $\alpha, \beta \in (0, 1)$ we have $f_\delta(\alpha) - f_\delta(\beta) \geq D$.

Proof. Let us calculate the Fourier transform of $f = f_\delta$ as 1-periodic function. Note that

$$\begin{aligned} \hat{f}(n) &= \int_{[0,1]} e^{-2\pi i n x} f(x) \, dx = \int_{[0,1]} e^{-2\pi i n x} \sum_{\lambda \in \mathbb{Z}} e^{-\frac{(x-\lambda)^2}{\delta}} \, dx \\ &= \int_{\mathbb{R}} e^{-2\pi i n x} e^{-\frac{x^2}{\delta}} \, dx = \sqrt{\pi \delta} e^{-\pi^2 n^2 \delta}, \end{aligned}$$

and therefore the function is not constant. Furthermore, by Parseval we have that

$$\int_{[0,1]} (f(x) - \hat{f}(0))^2 \, dx = \sum_{n=1}^{\infty} \hat{f}^2(n) = \sum_{n=1}^{\infty} \pi \delta e^{-2\pi n^2 \delta} \geq \pi \int_{\delta}^{\infty} e^{-2\pi \frac{x^2}{\delta}} \, dx$$

and so the variance of f is bounded below whenever δ is bounded, proving the result. \square

We now prove our main result, the failure of a local central limit theorem for constant p and k .

Theorem 1.1. *Fix $p \in (0, 1)$ and $k \geq 3$. Then for all sufficiently large n relatively prime to $(k-1)!$ there is a point x such that*

$$\left| \mathbb{P}[\mathbf{kAP}(z) = x] - \frac{1}{\sigma_n \sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n)^2}{2\sigma_n^2}\right) \right| = \Omega\left(\frac{1}{\sigma_n}\right),$$

where μ_n and σ_n in the statement are the expectation and standard deviation of $\mathbf{kAP}(z)$ and z is sampled i.i.d. with probability p .

Throughout this section we use x_i to denote the 0,1 indicator of whether i is in our random set, and y_i to denote the normalized Bernoulli random variables $y_i := (x_i - p)/\sqrt{p(1-p)}$. We additionally use the shorthand $\ell = \sum_{i=1}^n y_i$ and $\tilde{\ell} := \sum_{i=1}^n x_i$ and $q := 1 - p$.

First we need exact formulae for \mathbf{kAP}^1 and \mathbf{kAP}^2 in terms of ℓ , which are

$$\mathbf{kAP}^1 = \sum_{i=1}^n \frac{k(n-1)}{2} p^{k-\frac{1}{2}} q^{\frac{1}{2}} y_i = \frac{k(n-1)}{2} p^{k-\frac{1}{2}} q^{\frac{1}{2}} \ell$$

$$\mathbf{kAP}^2 = \binom{k}{2} p^{k-1} q \sum_{|S|=2} y_S = \binom{k}{2} \frac{p^{k-1} q}{2} \left(\ell^2 - \sum_{i=1}^n y_i^2 \right) = \binom{k}{2} \frac{p^{k-1} q}{2} \left(\ell^2 - n - \frac{1-2p}{\sqrt{pq}} \ell \right)$$

where $y_S = \prod_{i \in S} y_i$. And so we find that we can express $\mathbf{kAP}^1 + \mathbf{kAP}^2 = C_0 + C_1 \ell + C_2 \ell^2 := Q(\ell)$ where

$$\begin{aligned} C_0 &:= -n \frac{k(k-1)}{4} p^{k-1} q \\ C_1 &:= \frac{k(n-1)}{2} p^{k-\frac{1}{2}} q^{\frac{1}{2}} - \frac{(1-2p)k(k-1)}{4} p^{k-\frac{3}{2}} q^{\frac{1}{2}} \\ C_2 &:= \frac{k(k-1)}{4} p^{k-1} q \end{aligned}$$

So to understand $\mathbf{kAP}^1 + \mathbf{kAP}^2$ it is enough to understand ℓ . We note that ℓ is valued on the lattice $-n\sqrt{p/q} + \mathbb{Z}/\sqrt{pq}$. The most commonly taken value for ℓ occurs when $\tilde{\ell} = \sum_{i=1}^n x_i = [pn]$. When this occurs we see that ℓ takes the value

$$a_0 := \frac{\tilde{\ell} - pn}{\sqrt{pq}} = \frac{[pn] - pn}{\sqrt{pq}}$$

and hence $\mathbf{kAP}^1 + \mathbf{kAP}^2$ takes the value

$$x_0 := \mathbf{kAP}^1 + \mathbf{kAP}^2 = Q(a_0) = C_0 + C_1 a_0 + C_2 a_0^2$$

Now we can define $X = \mathbf{kAP}^1(y) + \mathbf{kAP}^2(y) - x_0$ and $Y = \mathbf{kAP}^{\geq 3}(y)$. So $X + Y = \mathbf{kAP} - \mu - x_0$, and the most likely value taken by X is 0.

Let $\dots, A_{-2}, A_{-1}, A_0 = 0, A_1, \dots$ be the values taken by X when $|\ell| \lesssim_{k,p} n$, listed in order. In general we see that X takes the value A_t whenever $\tilde{\ell} = [pn] + t$ and so $\ell = a_0 + t/\sqrt{pq}$ when $|\ell| \lesssim_{k,p} n$ (as $Q'(\ell) > 0$ in such a range). Therefore we can compute for any t that

$$A_t = Q\left(a_0 + \frac{t}{\sqrt{pq}}\right) - Q(a_0) = C_2 \left(\frac{2ta_0}{\sqrt{pq}} + \frac{t^2}{pq} \right) + C_1 \frac{t}{\sqrt{pq}}$$

To alleviate notation we define a constant G for the dominant increment $G := C_1/\sqrt{pq}$. It will commonly be helpful to have the bound

$$|A_t - tG| = \left| C_2 \left(\frac{2ta_0}{\sqrt{pq}} + \frac{t^2}{pq} \right) \right| \leq C_2 \frac{t^2 + 2t}{pq}.$$

Now we are in a position to define the events we look at. Fix $\alpha \in [0, 1]$, and $i, B \in \mathbb{R}$. Then we define the intervals $I_\alpha(i, B)$ and families of intervals $L_\alpha(B, s)$ by setting

$$\begin{aligned} I_\alpha(i, B) &:= [G(i + \alpha) - B, G(i + \alpha) + B] \\ L_\alpha(B, s) &:= \bigcup_{i=-s}^s I_\alpha(i, B) \end{aligned}$$

Note that if $\alpha = 0$ then the intervals $I_0(i, B)$ are just intervals of length $2B$ centered around that lattice points in $C_1 \mathbb{Z}/\sqrt{pq}$. These intervals are disjoint so long as we ensure $|B| < G/2$. As α moves between 0 and 1, the interval slides between neighboring lattice points in $G\mathbb{Z}$. $L_\alpha(B, s)$ collects the most central $2s + 1$ intervals in the collection.

Our goal becomes to show that for some pair $B, s = o(n)$ there exist distinct values $\alpha, \beta \in (0, 1)$ so that $\mathbb{P}[X + Y \in L_\alpha(B, s)]$ and $\mathbb{P}[X + Y \in L_\beta(B, s)]$ are far. This will contradict the existence of a local limit theorem, thus finishing our proof of Section 8.

We first choose η so that $\mathbb{P}[|Y| \geq \eta C_1/\sqrt{pq}] \leq n^{-100}$. By hypercontractivity concentration bounds (see e.g. [O'D14, Theorem 10.24]) we see that we may take $\eta = \Theta_{p,k}((\log n)^{k/2})$.

We may use Lemma 7.1 and some work to compute $\mathbb{P}[X + Y \in L_\alpha(B, s)]$. We state our result and postpone the calculation to Appendix A.

Lemma 8.2. *Assume that $0 < B < n^{1-1/36}$ and $\eta < s < n^{1/2-1/24}$. Let σ_Y denote the standard deviation of Y (and as such $\sigma_Y = \Theta(n)$). Then*

$$\mathbb{P}[X + Y \in L_\alpha(B, s)] = \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_\delta(\alpha) + O\left(\eta n^{-1/8}\right)$$

for some uniformly bounded $\delta = \delta(n, k, p)$.

Remark. In fact, as $n \rightarrow \infty$ this function tends towards a limit; we do not bother replacing $\delta(n, k, p)$ with its limit $\delta(k, p)$ as this will be inconsequential to our arguments.

Additionally, were **kAP** to obey a local limit theorem, then we could compute $\mathbb{P}[X + Y \in L_\alpha(B, s)]$ in a different way.

Lemma 8.3. *Assume that Z is a random variable with mean μ_Z and standard deviation σ_Z which for all $m \in \mathbb{N}$ satisfies*

$$\mathbb{P}[Z = m] = \frac{1}{\sqrt{2\pi}\sigma_Z} \exp\left(-\frac{(m - \mu_Z)^2}{2\sigma_Z^2}\right) + o(\sigma_Z^{-1})$$

Then for any set $S \subset \mathbb{Z} \cap [\mu - T, \mu + T]$ with $T \leq \sigma_Z$ we have

$$\mathbb{P}(Z \in S) = \frac{|S|}{\sqrt{2\pi}\sigma_Z} + o\left(\frac{|S|}{\sigma_Z}\right) + O\left(\frac{|S|T}{\sigma_Z^2}\right).$$

Proof. This follows from simply noting that $|1 - \exp((m - \mu)^2/2\sigma_Z^2)| \lesssim (m - \mu)/\sigma_Z$ in this range, and summing over $\sum_{m \in S} \mathbb{P}(Z = m)$. \square

We now have all of the tools to prove Theorem 1.1.

Proof of Theorem 1.1. Choose $\delta = \delta(k, p)$, which is constant. We know from Lemma 8.1 that there are $\alpha \neq \beta \in (0, 1)$ so that $f_\delta(\alpha) - f_\delta(\beta) = \Omega(1)$. If we set $B = \lfloor n^{1-1/36} \rfloor$ and $s = \lfloor n^{1/2-1/24} \rfloor$ then by Lemma 8.2 we have

$$\mathbb{P}[X + Y \in L_\alpha(B, s)] = \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_\delta(\alpha) + O\left(\eta n^{-1/8}\right)$$

and likewise for $L_\beta(B, s)$. Crucially, note that the main term has size $n^{-5/72}$ and so dominates the error term. Applying the same reasoning to $L_\beta(B, s)$ and taking differences yields

$$\begin{aligned} \mathbb{P}[X + Y \in L_\alpha(B, s)] - \mathbb{P}[X + Y \in L_\beta(B, s)] &= \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} [f_\delta(\alpha) - f_\delta(\beta)] + O\left(\eta n^{-1/8}\right) \\ &= \Omega(n^{-5/72}) \end{aligned}$$

However if we assume that **kAP** obeys a local central limit theorem, then so does $X + Y$. Note that $L_\alpha(B, s)$ consists of elements of size at most $Gs + B = O(n^{3/2-1/24})$, and $|L_\alpha(B, s)| = \Theta(n^{3/2-5/72})$. So by Lemma 8.3 we see that

$$\mathbb{P}[X + Y \in L_\alpha(B, s)] - \mathbb{P}[X + Y \in L_\beta(B, s)] = o\left(n^{-5/72}\right) + O\left(n^{-1/9}\right)$$

This is a direct contradiction. \square

9. CONCLUSION

We conclude by pointing out various open questions regarding the failure of the local central limit theorem. First, note that there is a gap between the theorems we proved and a total explanation of the behavior exhibited in Figure 1. An ideal theorem would prove more precisely that the distribution of $\mathbf{kAP}(S)$ tends to the convolution of two discrete Gaussians seemingly exhibited in the figure. Such a precise theorem would prove that $\mathbf{kAP}(S)$, conditional on the size of S , satisfies a local central limit theorem. Failing this precise local limit theorem, at least one could hope to understand how wildly the distribution oscillates in the following precise sense:

Question 9.1. Fix $\epsilon = 1/100$. Let S be a subset of $\mathbb{Z}/n\mathbb{Z}$ where each element is chosen independently with probability $1/2$ and set $\mu_n = \mathbb{E}[\mathbf{3AP}(S)]$. What is the largest constant C such that for infinitely many n there exist integers x_n, y_n such that $|x - \mu_n|, |y - \mu_n| \leq n^{\frac{3}{2}-\epsilon}$ which satisfy

$$\frac{\mathbb{P}[\mathbf{3AP}(S) = x_n]}{\mathbb{P}[\mathbf{3AP}(S) = y_n]} \geq C?$$

We argue in the subsection below that $C \approx 4.745$ works, however our method does suggest that in fact that C cannot be taken arbitrarily large. To be more precise, the largest constant C coming from our method is $\frac{\sup(f(x))}{\inf(f(x))}$ where $f(x) = \sum_{\lambda \in \mathbb{Z}} e^{-9(x-\lambda)^2}$. This is in fact the optimal constant if one can prove a sufficiently strong local limit theorem in the style of our above results. Furthermore, proving that C cannot be taken to be unbounded, along with the central limit theorem in the paper, would immediately show anti-concentration for the random variable $\mathbf{3AP}(S)$ at the correct level.

9.1. Computing C . Here we compute the value of C coming out of Theorem 1.1 by computing the resulting value of λ in the case $k = 3, p = \frac{1}{2}$. It is easy to show that

$$8 \cdot \mathbf{3AP}(y) = \frac{n(n-1)}{2} + \frac{3}{2}(n-1) \sum_{a \in \mathbb{Z}/n\mathbb{Z}} y_a + 3 \sum_{0 \leq a < b < n} y_a y_b + \sum_{a \in \mathbb{Z}/n\mathbb{Z}} \sum_{d \in [n/2]} y_a y_{a+d} y_{a+2d},$$

hence $\mathbf{kAP}^1(y)$ has steps of size $3n$ (since $y_a \in \{-1, 1\}$) while $\mathbf{kAP}^3(y)$ has mean zero and variance $\frac{n(n-1)}{2}$. Hence the associated normal has standard deviation of size $\frac{n}{\sqrt{2}}$. If we normalize $\mathbf{kAP}^1(y)$ to have steps of size 1, then $\mathbf{kAP}^3(y)$ will be normalized to have standard deviation $\frac{\sqrt{2}}{6}$ hence density proportional to e^{-9x^2} . Therefore we find value $\delta = \frac{1}{9}$, and the claimed ratio as above. It can further be shown that the minimum is attained at $\frac{1}{2} + \mathbb{Z}$ and the maximum at \mathbb{Z} , hence we can prove a ratio of

$$C = \frac{\sum_{x \in \mathbb{Z}} e^{-9x^2}}{\sum_{x \in \mathbb{Z}} e^{-9(x-\frac{1}{2})^2}} \approx 4.745.$$

REFERENCES

- [BAKL⁺19] Yacine Barhoumi-Andréani, Christoph Koch, Hong Liu, et al. Bivariate fluctuations for the number of arithmetic progressions in random sets. *Electronic Journal of Probability*, 24, 2019.
- [Bera] Ross Berkowitz. A local limit theorem for cliques in $G(n, p)$. arXiv:1811.03527.
- [Berb] Ross Berkowitz. A quantitative local limit theorem for triangles in random graphs. arXiv:1610.01281.
- [BGSZ20] Bhaswar B. Bhattacharya, Shirshendu Ganguly, Xuancheng Shao, and Yufei Zhao. Upper Tail Large Deviations for Arithmetic Progressions in a Random Set. *Int. Math. Res. Not. IMRN*, (1):167–213, 2020.
- [CCHT] Bryce Cai, Annie Chen, Ben Heller, and Eyob Tsegaye. Limit Theorems for Descents in Permutations and Arithmetic Progressions in $\mathbb{Z}/p\mathbb{Z}$. arXiv:1810.02425.
- [FKS] Jacob Fox, Matthew Kwan, and Lisa Sauermann. Anticoncentration for subgraph counts in random graphs. arXiv:1905.12749.
- [GK16] Justin Gilmer and Swastik Kopparty. A local central limit theorem for triangles in a random graph. *Random Structures Algorithms*, 48(4):732–750, 2016.
- [HMS] Matan Harel, Frank Mousset, and Wojciech Samotij. Upper tails via high moments and entropic stability. arXiv:1904.08212.

- [JW16] Svante Janson and Lutz Warnke. The lower tail: Poisson approximation revisited. *Random Structures Algorithms*, 48(2):219–246, 2016.
- [Mec09] Elizabeth Meckes. On Stein’s method for multivariate normal approximation. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. (IMS) Collect.*, pages 153–178. Inst. Math. Statist., Beachwood, OH, 2009.
- [O’D14] Ryan O’Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.
- [War17] Lutz Warnke. Upper tails for arithmetic progressions in random subsets. *Israel J. Math.*, 221(1):317–365, 2017.

APPENDIX A. PROOF OF LEMMA 8.2

We use all the same terminology and notation as in Section 8, including C_0, C_1, C_2 . First we need a lemma relating the probability that $X + Y \in L_\alpha(B, s)$ to a bound on the the joint distribution of X and Y .

Lemma A.1. *Assume that $s > \eta$ and $B - C_2 \frac{s^2+s}{pq} \geq 0$. Then*

$$\mathbb{P}[X + Y \in L_\alpha(B, s)] \geq \mathbb{P}\left[|X| \leq A_{s-\eta}, \quad Y \in L_\alpha(\eta, B - C_2 \frac{s^2+s}{pq})\right]$$

Proof. We show that, in fact, the event on the left hand side contains the event on the right. Let $X = A_t$ where $|t| \leq s - \eta$ and assume that $Y = G(i + \alpha) + b$ where $|i| \leq \eta$ and $b \leq B - C_2 \frac{s^2+s}{pq}$. Then we note that

$$\begin{aligned} |X + Y - G(t + i + \alpha)| &= \left| \left(tG + C_2 \left(\frac{2ta_0}{\sqrt{pq}} + \frac{t^2}{pq} \right) + G(i + \alpha) + b \right) - G(t + i + \alpha) \right| \\ &\leq C_2 \frac{s(s+1)}{pq} + b \leq B \end{aligned}$$

As $(t + i) \leq s$ it follows that if the event on the right hand side occurs, then $X + Y \in L_\alpha(B, s)$. \square

Lemma A.2. *Assume that $|B| < G/2$ and $C_2 \frac{(s+\eta+2)(s+\eta+3)}{pq} < G/2$. Then*

$$\mathbb{P}[X + Y \in L_\alpha(B, s)] \leq \mathbb{P}\left[|X| \leq A_{s+\eta+1}, \quad Y \in L_\alpha(\eta, B + C_2 \frac{(s+\eta)(s+\eta+1)}{pq})\right] + O(n^{-100})$$

Proof. First we note that

$$\mathbb{P}(X + Y \in L_\alpha(B, S)) \leq \mathbb{P}(X + Y \in L_\alpha(B, S) \text{ and } |Y| \leq \eta G) + O(n^{-100}).$$

So we will throughout condition on the event that $|Y| \leq \eta G$. In that event, we show that the event on the left hand side implies that the event on the right hand side occurs. First, we can simply upper bound $|X|$ by

$$|X| \leq |X + Y| + |Y| \leq |(s + \alpha)G + B| + \eta G.$$

But we know that

$$A_{s+\eta+2} \geq (s + \eta + 2)G - C_2 \frac{(s + \eta + 2)(s + \eta + 3)}{pq} > (s + \eta)G + \alpha G + B.$$

And so we have that $X = A_t$ where $|t| \leq s + \eta + 1$.

For Y we note that $X + Y \in L_\alpha(B, s)$ implies that $X + Y = (r + \alpha)G + b$ where $|r| \leq s$ and $|b| \leq B$. By the argument above we know that $X = A_t$ for $t \leq s + \eta + 1$. And so we can bound

$$\begin{aligned} |Y - (r - t + \alpha)G| &= |X + Y - X - (r - t + \alpha)G| = |(r + \alpha)G + b - A_t - (r - t + \alpha)G| \\ &\leq B + \left| C_2 \frac{(s + \eta + 1)(s + \eta + 2)}{pq} \right|. \end{aligned}$$

The last thing we need to do to show that $Y \in L_\alpha(\eta, B + C_2 \frac{(s+\eta)(s+\eta+1)}{pq})$ is to show that $|r - t| \leq \eta$, but were it otherwise the above equation implies that $|Y| > (\eta + 1)G - G$, contradicting our assumption that $|Y| \leq \eta G$. Therefore it must follow that $Y \in L_\alpha(\eta, B + C_2 \frac{(s+\eta)(s+\eta+1)}{pq})$. \square

The main upshot here is that for well chosen values of B and s , the probability bounds furnished by the above lemmas will be indistinguishable up to a margin of error. The last ingredient we need is an estimate for the probability the events $|X| \leq A_s$ and $Y \in L_\alpha(B, s)$.

Lemma A.3. *For any interval of the form $(T - B, T + B)$ we have*

$$\int_{T-B}^{T+B} e^{-t^2/2} dt = 2Be^{-T^2/2} + O(B^3)$$

Proof. This is just the midpoint rule combined with the observation that $|\frac{d^2}{dt^2} e^{-t^2/2}| \leq 1$. \square

Lemma A.4. *Assume that $B < n^{-1/36}$, $s < n^{1/2-1/24}$, and $\eta \geq \log(n)$. Let f_δ be as defined in Lemma 8.1. Then*

$$\mathbb{P}[|X| \leq A_s, Y \in L_\alpha(\eta, B)] = \frac{2sB\sqrt{2}}{\sigma_Y \sqrt{\pi npq}} f_\delta(\alpha) + O(\eta n^{-1/8})$$

where $\delta = pq\sigma_Y^2/C_1^2$.

Proof. First, we note that $|X| \leq A_s$ is equivalent to saying that $\overline{\mathbf{kAP}}^1 \in [a_0/\sqrt{n} - \frac{s}{\sqrt{npq}}, a_0/\sqrt{n} + \frac{s}{\sqrt{npq}}]$. First we note that by Lemma 7.1 for each interval $I_\alpha(i, B)$ we have

$$\mathbb{P}(|X| \leq A_s, Y \in I_\alpha(i, B)) = \left(\frac{1}{\sqrt{2\pi}} \int_{a_0/\sqrt{n} - \frac{s}{\sqrt{npq}}}^{a_0/\sqrt{n} + \frac{s}{\sqrt{npq}}} e^{-t^2/2} dt \right) \left(\frac{1}{\sqrt{2\pi}} \int_{I_\alpha(i, B)/\sigma_Y} e^{-t^2/2} dt \right) + O(n^{-1/8})$$

By Lemma A.3 we can estimate both of these integrals. The first is

$$\frac{1}{\sqrt{2\pi}} \int_{a_0/\sqrt{n} - \frac{s}{\sqrt{npq}}}^{a_0/\sqrt{n} + \frac{s}{\sqrt{npq}}} e^{-t^2/2} dt = \left(\frac{2s}{\sqrt{npq}} \right) \frac{1}{\sqrt{2\pi}} e^{-a_0^2/2\sigma_1^2} + O\left(\frac{s^3}{(npq)^{3/2}} \right) = \frac{s\sqrt{2}}{\sqrt{\pi npq}} + O\left(\frac{s^3}{n^{3/2}} + \frac{1}{n^3} \right)$$

and the second estimate is

$$\frac{1}{\sqrt{2\pi}} \int_{I_\alpha(i, B)/\sigma_Y} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_{\frac{C_1(i+\alpha)}{\sqrt{pq}\sigma_Y} - B/\sigma_Y}^{\frac{C_1(i+\alpha)}{\sqrt{pq}\sigma_Y} + B/\sigma_Y} e^{-t^2/2} dt = \frac{2B}{\sigma_Y} e^{-\frac{C_1^2(i+\alpha)^2}{pq\sigma_Y^2}} + O(B^3/n^3).$$

So, defining $\delta := pq\sigma_Y^2/C_1^2 = \Theta(1)$ and combining all of these estimates, we find that

$$\begin{aligned} \mathbb{P}[|X| \leq A_s, Y \in I_\alpha(i, B)] &= \left(\frac{s\sqrt{2}}{\sqrt{\pi npq}} + O\left(\frac{s^3}{n^{3/2}} + \frac{1}{n^3} \right) \right) \left(\frac{2B}{\sigma_Y} e^{-\frac{(i+\alpha)^2}{\delta}} + O(B^3/n^3) \right) + O(n^{-1/8}) \\ &= \frac{2sB\sqrt{2}}{\sigma_Y \sqrt{\pi npq}} e^{-\frac{(i+\alpha)^2}{\delta}} + O\left(s^3/n^{3/2} + \frac{B^3 s}{n^{3.5}} + n^{-1/8} \right) \\ &= \frac{2sB\sqrt{2}}{\sigma_Y \sqrt{\pi npq}} e^{-\frac{(i+\alpha)^2}{\delta}} + O\left(n^{-1/8} \right). \end{aligned}$$

Thus taking a union yields

$$\mathbb{P}[|X| \leq A_s, Y \in L_\alpha(\eta, B)] = \sum_{i=-\eta}^{\eta} \frac{2sB\sqrt{2}}{\sigma_Y \sqrt{\pi npq}} e^{-\frac{(i+\alpha)^2}{\delta}} + O(\eta n^{-1/8})$$

$$\begin{aligned}
&= \sum_{i=-\infty}^{\infty} \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} e^{-\frac{(i+\alpha)^2}{\delta}} + O\left(\eta n^{-1/8} + e^{-\eta^2/\delta}\right) \\
&= \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8}\right). \quad \square
\end{aligned}$$

Lemma 8.2. *Assume that $|B| < n^{1-1/36}$, $s < n^{1/2-1/24}$ and $\eta < s$. Let σ_Y denote the standard deviation of Y (and as such $\sigma_Y = \Theta(n)$). Then*

$$\mathbb{P}[X + Y \in L_{\alpha}(B, s)] = \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8}\right).$$

Proof. First we use Lemmas A.2 and A.4 to upper bound

$$\begin{aligned}
\mathbb{P}[X + Y \in L_{\alpha}(B, s)] &\leq \mathbb{P}\left[|X| \leq A_{s+\eta+1}, \quad Y \in L_{\alpha}\left(\eta, B + C_2 \frac{(s+\eta)(s+\eta+1)}{pq}\right)\right] + O(n^{-100}) \\
&= \frac{s\left(B + C_2 \frac{(s+\eta)(s+\eta+1)}{pq}\right) 2\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8}\right) \\
&= \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8} + s^3/n^{1.5}\right).
\end{aligned}$$

Next we use Lemma A.1 to lower bound

$$\begin{aligned}
\mathbb{P}[X + Y \in L_{\alpha}(B, s)] &\geq \mathbb{P}\left[|X| \leq A_{s-\eta}, \quad Y \in L_{\alpha}\left(\eta, B - C_2 \frac{s^2+s}{pq}\right)\right] \\
&= \frac{(s-\eta)\left(B - C_2 \frac{s^2+s}{pq}\right) 2\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8}\right) \\
&= \frac{2sB\sqrt{2}}{\sigma_Y\sqrt{\pi npq}} f_{\delta}(\alpha) + O\left(\eta n^{-1/8} + s^3/n^{1.5} + \eta B/n^{1.5}\right).
\end{aligned}$$

By our hypotheses on B , s , and η the $\eta n^{-1/8}$ term dominates in both inequalities. □

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06520, USA
E-mail address: ross.berkowitz@yale.edu

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
E-mail address: asah@mit.edu

MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
E-mail address: msawhney@mit.edu