

Bivariate fluctuations for the number of arithmetic progressions in random sets*

Yacine Barhoumi-Andréani[†] Christoph Koch[‡] Hong Liu[§]

Abstract

We study arithmetic progressions $\{a, a + b, a + 2b, \dots, a + (\ell - 1)b\}$, with $\ell \geq 3$, in random subsets of the initial segment of natural numbers $[n] := \{1, 2, \dots, n\}$. Given $p \in [0, 1]$ we denote by $[n]_p$ the random subset of $[n]$ which includes every number with probability p , independently of one another. The focus lies on sparse random subsets, i.e. when $p = p(n) = o(1)$ as $n \rightarrow +\infty$. Let X_ℓ denote the number of distinct arithmetic progressions of length ℓ which are contained in $[n]_p$. We determine the limiting distribution for X_ℓ not only for fixed $\ell \geq 3$ but also when $\ell = \ell(n) \rightarrow +\infty$ with $\ell = o(\log n)$. The main result concerns the joint distribution of the pair $(X_\ell, X_{\ell'})$, $\ell > \ell'$, for which we prove a bivariate central limit theorem for a wide range of p . Interestingly, the question of whether the limiting distribution is trivial, degenerate, or non-trivial is characterised by the asymptotic behaviour (as $n \rightarrow +\infty$) of the threshold function $\psi_\ell = \psi_\ell(n) := np^{\ell-1}\ell$. The proofs are based on the method of moments and combinatorial arguments, such as an algorithmic enumeration of collections of arithmetic progressions.

Keywords: arithmetic progression; central limit theorem; bivariate fluctuations; method of moments; exploration process.

AMS MSC 2010: 60C05; 11B25; 05C80; 60F05.

Submitted to EJP on April 2, 2019, final version accepted on November 7, 2019.

Supersedes arXiv:1902.04176.

*Y.B. was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant No. EP/L012154/1. C.K. was supported by Engineering and Physical Sciences Research Council (EPSRC) Grant No. EP/N004833/1 and the European Research Council (ERC) Grant No. 639046. H.L. was supported by the UK Research and Innovation Future Leaders Fellowship MR/S016325/1, and the Leverhulme Trust Early Career Fellowship ECF-2016-523.

[†]Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstrasse 150, 44780 Bochum, Germany.
E-mail: yacine.barhoumi@rub.de

[‡]Department of Statistics, University of Oxford, St. Giles 24-29, Oxford OX1 3LB, UK.
E-mail: christoph.koch@stats.ox.ac.uk

[§]Mathematics Institute and DIMAP, University of Warwick, Coventry, CV4 7AL, UK.
E-mail: h.liu.9@warwick.ac.uk

1 Introduction and main results

An ℓ -term arithmetic progression (ℓ -AP) in a set $\mathcal{X} \subset \mathbb{Z}$ is an (ordered) ℓ -tuple of distinct numbers $(a, a + b, \dots, a + (\ell - 1)b)$ whose elements belong to \mathcal{X} . In Dickson's *History of the Theory of Numbers*, the analysis of APs is traced back to around 1770 when it became prominent due to Lagrange and Waring investigating how large the common difference of an ℓ -AP of primes must be. Ever since, the study of APs has remained an extremely active domain of research and led to several results of fundamental importance, for instance Dirichlet's Theorem [12] proved in 1837 played a key role in the formation of analytic number theory. Perhaps unsurprisingly, APs also became objects of interest in other fields such as combinatorics: van der Waerden's celebrated theorem [36] states that for any given positive integers r and k , there exists some number $W(r, k)$ (the minimal such number being nowadays called the van der Waerden number) such that if the integers $\{1, 2, \dots, W(r, k)\}$ are colored with one of r different colors, then there exist at least k integers in arithmetic progression whose elements are of the same color.

Erdős also stated a number of conjectures related to ℓ -APs [5, pp. 232-233]. In particular, he offered \$1000 to solve the following *largest progression-free subset* problem: find the cardinality of the largest subset of $\{1, \dots, m\}$ ($m \in \mathbb{N}$) which does not contain any ℓ -AP. This problem was solved by Szémeredi with his celebrated density theorem [35]: a subset of \mathbb{N} of non-zero upper asymptotic density contains ℓ -APs of any arbitrary length ℓ . The case $\ell = 3$ was settled in Roth's celebrated theorem that opened the use of Fourier analysis in additive combinatorics. Subsequently, based on Szémeredi's Theorem, Green and Tao [18] proved the long-standing conjecture on prime APs: (dense subsets of) the primes contain infinitely many ℓ -APs for all lengths ℓ .

In 1936, Cramér [11] conjectured that the gaps between two consecutive primes remain asymptotically bounded by the square of their logarithms and backed this conjecture with a heuristic model that replaces the set \mathcal{P} of primes by a random set \mathcal{P}' made out of Bernoulli random variables, where $\mathbb{P}(m \in \mathcal{P}') \approx 1/\log m$ independently for all integers $m \geq 2$. However, the study of APs in random sets does not only provide a nice heuristic for number theoretic problems but is also a very natural and interesting model from a probabilistic point of view. For instance, Kohayakawa, Łuczak, and Rödl [24] proved that *sparse* uniformly random subsets $M \subseteq \{1, \dots, n\}$ of size $|M| = \Omega(\sqrt{n})$ have the property that any (sufficiently) dense subset of M already contains a 3-AP with probability tending to 1 as $n \rightarrow +\infty$. For recent developments on extremal theorems for random sets (not only ℓ -APs), we refer to [2, 10, 33, 34].

In this article we focus our attention on longer APs in sparse binomial subsets of $\{1, \dots, n\}$, including ℓ -APs with length $\ell = \ell(n) \rightarrow +\infty$ as $n \rightarrow +\infty$. In particular, we determine the limiting distribution of the number of ℓ -APs and analyse the joint distribution of the numbers of ℓ -APs and ℓ' -APs of different lengths $\ell \neq \ell'$.

1.1 Main results

We consider a family of random subsets of the initial segments $[n] := \{1, \dots, n\} \subset \mathbb{N}$ of the integers. For any $p = p(n) \in [0, 1]$ let Ξ_1, \dots, Ξ_n be a collection of independent identically distributed $\text{Be}(p)$ random variables, denote their product measure by \mathbb{P} , and let $[n]_p := \{i \in [n] : \Xi_i = 1\}$ be the p -percolation of $[n]$, i.e. $[n]_p$ is the random subset of $[n]$ obtained by deleting any of the elements with probability $1 - p$, independently of all other elements. We use the term *constant* to mean independent of the parameter n , and any unspecified asymptotic notation (including limits) is to be understood with respect to $n \rightarrow +\infty$.

For any integer $\ell \in \{3, \dots, n\}$ we denote the set of all ℓ -APs in $[n]$ by \mathcal{A}_ℓ and define X_ℓ to be the random variable counting the number of ℓ -APs in $[n]_p$, namely

$$X_\ell = X_\ell(n) := \sum_{T \in \mathcal{A}_\ell} \mathbb{1}_{\{T \subseteq [n]_p\}}.$$

Clearly, $[n]$ itself is an n -AP and any ℓ -AP contains a whole number of ℓ' -APs for each $3 \leq \ell' \leq \ell - 1$. Therefore, the family $\{X_\ell\}_{3 \leq \ell \leq n}$ is obviously correlated in a non-trivial way. While the FKG inequality (see Theorem 2.10) implies that this family is actually positively correlated, it is a priori unclear whether this correlation is asymptotically relevant. The main goal of this article is to study the asymptotic behaviour of the joint distribution of the pair (X_{ℓ_1}, X_{ℓ_2}) with $\ell_1 > \ell_2$.

We start by determining the limiting distribution of the number of ℓ -APs to be either a Poisson distribution or a Gaussian distribution. Let $\sigma_\ell := \sqrt{\mathbb{V}(X_\ell)}$ denote the standard deviation of X_ℓ .

Theorem 1.1 (Univariate limiting distributions). *Let $\ell \geq 3$ be either a constant, or $\ell = \ell(n) \rightarrow +\infty$ satisfying $\ell/\log n \rightarrow 0$, and let $0 < p = p(n) = o(1)$.*

(a) *If $n^2 p^\ell / (\ell - 1) \rightarrow c$, for some $c \in \mathbb{R}_+$, then $X_\ell \xrightarrow{d} \text{Po}(c/2)$.*

(b) *If $n^2 p^\ell / (\ell - 1) \rightarrow +\infty$, then $(X_\ell - \mathbb{E}(X_\ell)) \sigma_\ell^{-1} \xrightarrow{d} \text{N}(0, 1)$.*

While a priori ℓ could be as large as n , it is easy to see that the random subset $[n]_p$ with $p = o(1)$ (i.e. in the sparse regime) asymptotically almost surely (a.a.s.) does not contain any ℓ -APs with $\ell = \ell(n) \geq C \log n$ for any constant $C > 0$. This follows by a first moment argument, since

$$\mathbb{E}(X_\ell) \stackrel{Cl. 2.1}{=} (1 \pm o(1)) \frac{n^2 p^\ell}{2(\ell - 1)} \leq \exp [2 \log n - C \log n \log(p^{-1})] = o(1), \quad (1.1)$$

and thus by Markov's inequality $\mathbb{P}(X_\ell = 0) \rightarrow 1$. In other words, Theorem 1.1 is optimal concerning the range of ℓ .¹

Theorem 1.1 shares conceptual similarities with a result of Ruciński [30] that deals with the number of copies of a given graph H in a binomial random graph $G(n, p)$ obtained as the p -percolation of the complete graph K_n . While [30, Thm. 2] deals only with graphs H of fixed size, it considers any possible graph having at least one edge; moreover, the method employed to prove it is the method of moments that will be used in the proof of Theorem 1.2 but that could also be used for Theorem 1.1.

We remark that for constant $\ell \geq 3$, Theorem 1.1 hardly comes as a surprise since X_ℓ is a sum of “weakly dependent” Bernoulli random variables. The Gaussian approximation follows then from a sufficient criterion due to Mikhailov (cf. Theorem 3.5), while the Chen-Stein method (cf. Theorem 3.2) yields the Poisson approximation. Note that in the case where ℓ is not allowed to grow with n , the Poisson regime is in fact a consequence of [31, Thm. 1.3] that deals with more general systems of equations in random sets, and also of [30, Thm. 1] where the notion of Poisson-convergence, due to Barbour, is used, this notion implying in particular convergence in distribution to a Poisson random variable. Yet, in the general case that we treat, we could not find a full proof of this result in the literature. The fact that the proof carries through for growing $\ell = \ell(n) \rightarrow +\infty$ is largely due to the fact that the expectation in (1.1) decreases exponentially quickly in ℓ .

Our main result characterises the bivariate fluctuations of the pair (X_{ℓ_1}, X_{ℓ_2}) when both random variables are within their respective Gaussian regimes, as determined in Theorem 1.1.

¹Except for cases where we can only expect convergence along subsequence, for instance if $\ell = \ell(n)$ alternates (periodically) between two or more constants.

Theorem 1.2 (Bivariate fluctuations for APs of different lengths). *For $i \in \{1, 2\}$, let $\ell_i \geq 3$ be either a constant, or $\ell_i = \ell_i(n) \rightarrow +\infty$, such that we have $\ell_2 < \ell_1$ point-wise² and $\ell_1/\log n \rightarrow 0$. Let $0 < p = p(n) < 1$ be such that $p\ell_1^9 \rightarrow 0$ and $n^2p^{\ell_1}\ell_1^{-9} \rightarrow +\infty$. Then we have*

$$\left(\frac{X_{\ell_1} - \mathbb{E}(X_{\ell_1})}{\sigma_{\ell_1}}, \frac{X_{\ell_2} - \mathbb{E}(X_{\ell_2})}{\sigma_{\ell_2}} \right) \xrightarrow{d} \mathbb{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \kappa_{\ell_1, \ell_2} \\ \kappa_{\ell_1, \ell_2} & 1 \end{pmatrix} \right),$$

where κ_{ℓ_1, ℓ_2} satisfies

$$\begin{cases} \kappa_{\ell_1, \ell_2} = 0, & \text{if } np^{\ell_1-1}\ell_1 \rightarrow 0; \\ 0 < \kappa_{\ell_1, \ell_2} < 1, & \text{if } np^{\ell_1-1}\ell_1 \rightarrow c \in \mathbb{R}_+ \vee [np^{\ell_1-1}\ell_1 \rightarrow +\infty \wedge \ell_2 \text{ is a constant}]; \\ \kappa_{\ell_1, \ell_2} = 1, & \text{if } np^{\ell_1-1}\ell_1 \rightarrow +\infty \wedge \ell_2 = \ell_2(n) \rightarrow +\infty. \end{cases}$$

Interestingly, the strength of the correlation is characterised by the asymptotic behaviour of the function

$$\psi_{\ell_1} = \psi_{\ell_1}(n) := np^{\ell_1-1}\ell_1, \tag{1.2}$$

which originates from the combinatorial structure of tuples of overlapping APs. There are two structures, *loose pairs* and *overlap pairs* (see Definition 2.2), which compete to dominate the centralised second moments of the pair (X_{ℓ_1}, X_{ℓ_2}) . The function ψ_{ℓ_1} is obtained as the ratio of the contribution of loose pairs by that of overlap pairs (of ℓ_1 -APs); when $\psi_{\ell_1} \rightarrow 0$, overlap pairs dominate, and when $\psi_{\ell_1} \rightarrow +\infty$, loose pairs dominate. We call the former the *overlap pair regime*, and the latter the *loose pair regime*. An explicit expression of κ_{ℓ_1, ℓ_2} is given in Lemma 2.11 and its proof; its derivation is surprisingly quite intricate and involves an integral representation.

Furthermore, we want to highlight that when $\ell_2 = \ell_2(n) \rightarrow +\infty$ (and thus also $\ell_1 = \ell_1(n) \rightarrow +\infty$), the random variables X_{ℓ_1} and X_{ℓ_2} are either asymptotically uncorrelated, or converge to the *same* random variable (once renormalised). However, in all other cases, there exists a regime where the asymptotic correlation is non-trivial.

Lastly, we remark that the slightly more restrictive conditions $p\ell_1^9 \rightarrow 0$ and $n^2p^{\ell_1}\ell_1^{-9} \rightarrow +\infty$ are an artefact of the proof method, we strongly believe that the result remains true under the weaker assumptions $p \rightarrow 0$ and $n^2p^{\ell_1-1}\ell_1^{-1} \rightarrow +\infty$, which characterise the sparse Gaussian regime for ℓ_1 -APs, cf. Theorem 1.1(b).

1.2 Related work

In the literature, the study of X_ℓ for random subsets of the integers is largely focused on $\ell \geq 3$ being a constant and estimating the probability of large deviations from its mean, i.e. the *upper tail* probabilities $\mathbb{P}(X_\ell \geq (1 + \varepsilon)\mathbb{E}(X_\ell))$, and the *lower tail* probabilities $\mathbb{P}(X_\ell \leq (1 - \varepsilon)\mathbb{E}(X_\ell))$. For a recent survey on large deviations in random graphs (and related combinatorial structures) see [8].

For the upper tail, Janson and Ruciński [22] obtained upper and lower bounds on $-\log \mathbb{P}(X_\ell \geq (1 + \varepsilon)\mathbb{E}(X_\ell))$ being apart by a factor of $\log(1/p)$ by extending an earlier result by Janson, Oleszkiewicz, and Ruciński [21] on large deviations for subgraph counts in random graphs. Subsequently, Warnke [37] closed this gap by proving that

$$-\log \mathbb{P}(X_\ell \geq (1 + \varepsilon)\mathbb{E}(X_\ell)) = \Theta_\varepsilon(\Phi(\mathbb{E}(X_\ell))), \quad \Phi(x) := \min\{x, \sqrt{x} \log(1/p)\},$$

and also supplying the dependency on ε of the implied constants in Θ_ε . Notably, provided that p is in the loose pair regime (more precisely, $\psi_\ell \geq \log n$, where $\psi_\ell = n\ell p^{\ell-1}$ as in (1.2)) the results in [37] also extend to *moderate variations*, i.e. events of the form $\{X_\ell \geq \mathbb{E}(X_\ell) + t\}$ for any $t \geq \sigma_\ell$. Complementing these results, Bhattacharya, Ganguly,

²I.e. $\ell_2(n) < \ell_1(n)$ for all $n \geq 1$.

Shao, and Zhao [3] pinned down the precise large deviation rate function for “sufficiently large” p . By contrast to the approach in [37], the proof in [3] builds on the non-linear large deviation principle by Chatterjee and Dembo [9] and its refinement due to Eldan [13] in terms of the concept of *Gaussian width*, a particular notion of complexity. Recently, Briët and Gopi [7] derived an upper bound on the Gaussian width leading to an improvement of the lower bound on p given in [3]. The special case $\ell = 3$ was already included in [9].

On the other hand, the lower tail has received less attention: for all constants $\ell \geq 3$, Janson and Warnke [23] determined the large deviation rate function up to constants to be

$$-\log \mathbb{P}(X_\ell \geq (1 - \varepsilon)\mathbb{E}(X_\ell)) = \Theta(\varepsilon^2 \min\{\mathbb{E}(X_\ell), np\}),$$

while Mousset, Noever, Panagiotou, and Samotij [28] concentrated on the probability of $[n]_p$ to be ℓ -AP free, and expressed $-\log \mathbb{P}(X_\ell = 0)$ as an alternating sum of certain *joint cumulants* defined in terms of the *dependency graph* associated to X_ℓ . The results on ℓ -APs in [28] hold only for p within the overlap pair regime ($\psi_\ell = o(1)$, where $\psi_\ell = n\ell p^{\ell-1}$ as in (1.2)).

We complement the literature results on large and moderate deviations by considering typical deviations and thereby determining the limiting distribution of X_ℓ not only for all constants $\ell \geq 3$ but also when $\ell = \ell(n) \rightarrow +\infty$. Additionally, we also investigate the interaction of the number of APs of different length occurring in $[n]_p$, i.e. typical fluctuations of the pair $(X_\ell, X_{\ell'})$. Strikingly, we find a significantly different behaviour of their bivariate fluctuations in the overlap pair regime, as compared to the loose pair regime. By contrast to the results on moderate deviations in [37] or the result in [28] which work only in one of the two regimes, we employ the same approach in both regimes.

1.3 Proof method and outline

The main goal of this article lies in the analysis of bivariate fluctuations of the pair (X_{ℓ_1}, X_{ℓ_2}) based on the method of moments: we show that the joint moments of (X_{ℓ_1}, X_{ℓ_2}) , once centered and rescaled, converge to the moments of a Gaussian vector, which ensures the convergence in distribution. More formally, we apply the combination of the following two classical results:

Theorem 1.3 (e.g. Theorem 30.2 in [4]). *Let \mathcal{Y} be a random variable which is determined by its moments, and let $(Y_n)_{n \in \mathbb{N}}$ be a sequence of random variables having finite moments of all orders. If $\lim_{n \rightarrow +\infty} \mathbb{E}(Y_n^k) = \mathbb{E}(\mathcal{Y}^k)$ for all $k \in \mathbb{N}$, then $Y_n \xrightarrow[n \rightarrow +\infty]{d} \mathcal{Y}$.*

The same principle transfers to multivariate random variables, by application of the Cramér-Wold device.

Theorem 1.4 (Cramér-Wold device, e.g. Theorem 29.4 in [4]). *For any $r \in \mathbb{N}$, let $\mathcal{Y} = (\mathcal{Y}_1, \dots, \mathcal{Y}_r)$ and $Y_n = (Y_{n,1}, \dots, Y_{n,r})$, $n \in \mathbb{N}$, be random vectors. Then $Y_n \xrightarrow{d} \mathcal{Y}$ if and only if*

$$\sum_{i=1}^r u_i Y_{n,i} \xrightarrow{d} \sum_{i=1}^r u_i \mathcal{Y}_i, \quad \forall u_1, \dots, u_r \in \mathbb{R}.$$

Instances of this method in the setting of probabilistic combinatorics could not be traced back, but one can already see it in the proof by Füredi and Komlós [15] of the Wigner semi-circle distribution, or in the work of Ruciński [30], itself inspired by Maheara [25]. Our approach for the analysis of the (normalised) joint moments with growing ℓ was in particular inspired by a recent result of Gao and Sato [17] determining the limiting distribution of the number of matchings of size $\ell = \ell(n)$ in $G(n, p)$ to be either a Normal or

a Log-normal distribution. The scheme of the method consists in finding the appropriate combinatorial structure that describes the moments of the limiting distribution and to show that such a structure governs the moments of the actual (normalised) random variables under study. The method can also be extended to joint moments, as long as a relevant combinatorial structure underlies them. It is well-known that the odd moments of a centered, multivariate Gaussian distribution vanish, while the even moments can be expressed combinatorially: for $k \in \mathbb{N}$ the $2k$ -th moment is given by a sum over all perfect matchings of the set $[2k]$. The key technical point in [15, 17, 30] is thus to find a suitable *coding* of the moments that highlights the combinatorial structure giving the main contribution, hence, the heart of our proof lies in showing that the (even and centred) joint moments of (X_{ℓ_1}, X_{ℓ_2}) are dominated by a *matching structure*.

In fact, we will see that this combinatorial structure is encoded in the *dependency graph* Γ (cf. Definition 3.1) associated with the pair (X_{ℓ_1}, X_{ℓ_2}) . Depending on the range of p , the main contribution will come from matchings consisting of overlap pairs and/or loose pairs, and can be determined explicitly. It then remains to bound the contributions of all non-matching configurations. This last step is based on an algorithmic exploration of the components in Γ ; a similar argument was previously used by Bollobás, Cooley, Kang, and the second author [6] in the context of jigsaw percolation on random hypergraphs. By contrast, in [17] this last step was based on the *switching method* introduced by McKay [26], which turned out to be difficult to apply in the setting of APs due to their arithmetic structure.

We close with an outline of the article: Section 2 focusses on counting APs and pairs of APs, and deriving the joint second moments from these. Since we require a high level of precision, the counting argument for loose pairs of APs turns out to be surprisingly challenging. In Section 3 we complete the proof of Theorem 1.1 based on two sufficient criteria from the literature. The higher joint moments of the pair (X_{ℓ_1}, X_{ℓ_2}) are analysed in Section 4, where we also complete the proof of Theorem 1.2 and provide an alternative proof of Theorem 1.1(b). We then conclude with a discussion of open problems in Section 5.

2 Preliminaries: counting APs and pairs of APs

We start out with determining the asymptotics related to the set of APs in $[n]$. First, we consider the total number of ℓ -APs, denoted by $A_\ell := |\mathcal{A}_\ell|$, where we recall that \mathcal{A}_ℓ denotes the set of ℓ -APs in $[n]$.

Claim 2.1. *For any $3 \leq \ell = \ell(n) \leq n$, we have*

$$A_\ell = \begin{cases} (1 \pm o(1)) \frac{n^2}{2^{\ell-1}} & \text{if } \ell/n \rightarrow 0, \\ \Theta(n) & \text{if } \ell/n \rightarrow c \in (0, 1), \\ (1 \pm o(1))(n - \ell + 1) & \text{if } \ell/n \rightarrow 1. \end{cases}$$

In particular, the following asymptotics holds for all $3 \leq \ell = \ell(n) \leq n$:

$$A_\ell = \Theta(n(n - \ell + 1)\ell^{-1}).$$

Furthermore, for any $3 \leq \ell = \ell(n) = o(n)$, we have

$$\mathbb{E}(X_\ell) = A_\ell p^\ell = (1 \pm o(1)) \frac{n^2 p^\ell}{2^{\ell-1}}.$$

Proof. Let $R := \left(\frac{n-1}{\ell-1} - \left\lfloor \frac{n-1}{\ell-1} \right\rfloor \right) \cdot (\ell - 1)$ and observe that $0 \leq R \leq \ell - 2$. We have

$$\begin{aligned} A_\ell &= \sum_{\delta=1}^{\lfloor \frac{n-1}{\ell-1} \rfloor} \sum_{m=1}^n \mathbb{1}_{\{m+(\ell-1)\delta \leq n\}} = \sum_{\delta=1}^{\lfloor \frac{n-1}{\ell-1} \rfloor} (n - \delta(\ell - 1)) \\ &= \left\lfloor \frac{n-1}{\ell-1} \right\rfloor \cdot n - (\ell - 1) \left(\left\lfloor \frac{n-1}{\ell-1} \right\rfloor + 1 \right) = \frac{n(n - \ell + 1)}{2(\ell - 1)} + f(R, \ell), \end{aligned}$$

where $f(R, \ell) := \frac{(R+1)(\ell-1)-(R+1)^2}{2(\ell-1)}$. Furthermore, we observe that for all ℓ we have $0 \leq f(R, \ell) \leq (\ell - 1)/8$. It remains to distinguish three cases:

- if $\ell/n \rightarrow 0$, then $f(R, \ell) = o(n) = o(n^2/\ell)$ and the claim follows immediately,
- if $\ell/n \rightarrow c$ for some constant $c \in (0, 1)$, then $f(R, \ell) = O(n)$ and again the claim follows immediately,
- if $\ell/n \rightarrow 1$, the ℓ -AP contained in $[n]$ is clearly an interval, hence the number of such choices is $n - \ell + 1$, completing the proof. \square

2.1 Loose pairs and overlap pairs

Next, we consider pairs of APs of potentially different lengths, and distinguish them by the size of their intersection.

Definition 2.2. Let $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) \leq n$.

(a) For any $r \in [\ell']$, we define

$$\mathcal{D}_{\ell, \ell'}^{(r)} := \{(T, T') \in \mathcal{A}_\ell \times \mathcal{A}_{\ell'} : |T \cap T'| = r\}$$

to be the set of (ordered) pairs of APs intersecting in precisely r elements.

(b) We say that a pair $(T, T') \in \mathcal{A}_\ell \times \mathcal{A}_{\ell'}$ is a loose pair if $|T \cap T'| = 1$. We use the shorthand $\mathcal{B}_{\ell, \ell'} := \mathcal{D}_{\ell, \ell'}^{(1)}$ for the set of all loose pairs.

(c) We say that a pair $(T, T') \in \mathcal{A}_\ell \times \mathcal{A}_{\ell'}$ is an overlap pair if $|T \cap T'| = \ell'$, or equivalently $T' \subseteq T$. We use the shorthand $\mathcal{C}_{\ell, \ell'} := \mathcal{D}_{\ell, \ell'}^{(\ell')}$ for the set of all overlap pairs.

(d) We denote the cardinalities of these sets by $D_{\ell, \ell'}^{(\cdot)} := |\mathcal{D}_{\ell, \ell'}^{(\cdot)}|$, $B_{\ell, \ell'} := |\mathcal{B}_{\ell, \ell'}|$, and $C_{\ell, \ell'} := |\mathcal{C}_{\ell, \ell'}|$, respectively. Furthermore, whenever $\ell = \ell'$ we drop one of the lower indices, e.g. we use $D_\ell^{(2)} := D_{\ell, \ell}^{(2)}$.

Notice that $C_{\ell, \ell} = A_\ell$. Computing the asymptotic behaviour of the number of overlap pairs is a corollary of Claim 2.1.

Corollary 2.3. For all $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) = o(n)$ we have

$$C_{\ell, \ell'} = \Theta(1) \cdot n^2(\ell - \ell' + 1)/\ell'.$$

Proof. Note that the number of overlap pairs $(T_1, T_2) \in \mathcal{C}_{\ell, \ell'}$ is equal to $A_\ell \cdot M$, where M is the number of ℓ' -APs in $[\ell]$. Indeed, by Claim 2.1, we have $M = \Theta(\ell(\ell - \ell' + 1)/\ell')$ and $A_\ell = \Theta(n^2\ell^{-1})$ and the statement follows. \square

Similarly, we obtain an upper bound on the number of pairs intersecting in precisely r elements for $2 \leq r \leq \ell' - 1$. Despite being somewhat crude, this bound will suffice for our purposes.

Claim 2.4. For any $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) = o(n)$ and $2 \leq r \leq \ell' - 1$ we have

$$D_{\ell, \ell'}^{(r)} = O(n^2 \ell (\ell')^2).$$

Furthermore, in case $r \geq \lfloor 2\ell'/3 \rfloor + 1$, we have

$$D_{\ell, \ell'}^{(r)} = O(n^2 (\ell - r + 1) (\ell' - r + 1) / \ell').$$

Proof. Note that a pair $(T, T') \in \mathcal{D}_{\ell, \ell'}^{(r)}$ is already uniquely determined by choosing the first AP T , for which there are at most $O(n^2 \ell^{-1})$ many choices by Claim 2.1; and then fixing the relative position of the first two intersection elements within T and T' , for which there are at most ℓ^2 and $(\ell')^2$ many choices, respectively. The first claim follows by multiplying.

As for the second bound, assume that $r \geq 2\ell'/3$, then any pair $(T, T') \in \mathcal{D}_{\ell, \ell'}^{(r)}$ induces an overlap pair consisting of the ℓ -AP T and the r -AP $T \cap T'$. By definition the number of such pairs is $C_{\ell, r}$ and thus at most $O(n^2 (\ell - r + 1) / \ell')$, by Corollary 2.3. Next, observe that once T and $T \cap T'$ are chosen, the common difference of T' needs to be a divisor of the common difference of $T \cap T'$. However, since $r \geq \lfloor 2\ell'/3 \rfloor + 1$ we have $|T' \setminus T| \leq \ell' - \lfloor 2\ell'/3 \rfloor - 1 \leq \ell'/3 < r - 1$, implying that both $T \cap T'$ and T' have the same common difference. So we may only choose how many elements of $T' \setminus T$ are smaller than the smallest element of $T \cap T'$, the number of choices is at most $\ell' - r + 1$. Hence in total we obtain the claimed upper bound. \square

By contrast, determining the asymptotics of the number of loose pairs is much more difficult. In the following we will use the convention that $1/0 = +\infty$, $\min\{x, +\infty\} = x$, and $\bar{x} := 1 - x$ for all $x \in [0, 1]$. Moreover, for any $3 \leq \ell = \ell(n) \leq n$ we define a function μ_ℓ by setting

$$\mu_\ell(x) := \frac{1}{\ell - 1} \sum_{\iota=1}^{\ell} \mathbb{1}_{\{x \geq (\iota-1)/(\ell-1)\}}, \tag{2.1}$$

for all $x \in [0, 1]$. Furthermore, we define functions $h_\ell: [0, 1] \mapsto [0, 1]$ by the following Lebesgue-Stieltjes integral

$$h_\ell(x) := \int_0^1 \min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\} d\mu_\ell(a). \tag{2.2}$$

We start by proving two technical properties of these functions

Claim 2.5. For any constant $\ell \geq 3$ the function h_ℓ is non-negative and has the following properties:

(a) Uniformly for all $1/3 \leq x \leq 2/3$, we have

$$h_\ell(x) \geq \frac{1}{2(\ell - 1)}. \tag{2.3}$$

(b) For all $0 \leq x \leq \frac{1}{2(\ell-1)}$ we have

$$h_\ell(x) = \frac{1}{\ell - 1} + xH_{\ell-2}, \tag{2.4}$$

where $H_t := \sum_{j=1}^t 1/j$ denotes the t -th harmonic number.

Proof. For the first claim, we note that $\min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\} \geq 1/2$ for all $1/3 \leq a \leq 2/3$ and $1/3 \leq x \leq 2/3$. We conclude by noting that there is at least one ι in $\{1, 2, \dots, \ell\}$ such that $1/3 \leq (\iota - 1)/(\ell - 1) \leq 2/3$.

For the second claim, let $x \leq \frac{1}{2(\ell-1)}$ and note that for all $1 \leq \iota \leq \ell - 1$ we have $1 - \frac{\iota-1}{\ell-1} \geq \frac{1}{\ell-1} > x$ implying that

$$\min\left\{\frac{x}{1 - \frac{\iota-1}{\ell-1}}, \frac{1-x}{\frac{\iota-1}{\ell-1}}\right\} = \frac{x(\ell-1)}{\ell-\iota}.$$

Therefore, we obtain

$$h_\ell(x) = \frac{x + (1-x)}{\ell-1} + x \sum_{\iota=2}^{\ell-1} \frac{1}{\ell-\iota} = \frac{1}{\ell-1} + xH_{\ell-2},$$

as claimed. □

Next, let the (binary) entropy function $h_\infty : [0, 1] \mapsto [0, 1]$ be defined by

$$h_\infty(x) := \begin{cases} x \log(1/x) + \bar{x} \log(1/\bar{x}) & \text{if } 0 < x < 1, \\ 0 & \text{if } x = 0 \vee x = 1, \end{cases} \tag{2.5}$$

and observe that h_∞ is continuous on $[0, 1]$. The next statement shows that h_∞ is obtained naturally from h_ℓ when $\ell = \ell(n) \rightarrow +\infty$.

Claim 2.6. *For any $\ell = \ell(n) \rightarrow +\infty$ with $\ell = o(n)$, the function h_ℓ converges to h_∞ in L^2 as $n \rightarrow +\infty$.*

Proof. We first observe that $\{d\mu_\ell\}_{\ell \in \mathbb{N}}$ converges weakly to the uniform measure on $[0, 1]$ as $n \rightarrow +\infty$. Furthermore, the function $a \mapsto \min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\}$ is bounded and continuous for all $x \in [0, 1]$, and thus we have

$$h_\ell(x) = \int_0^1 \min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\} d\mu_\ell(a) = (1 \pm o(1)) \int_0^1 \min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\} da.$$

Moreover, for all $x \in (0, 1)$ we have

$$\begin{aligned} \int_0^1 \min\left\{\frac{x}{a}, \frac{\bar{x}}{a}\right\} da &= \int_0^1 \left(\frac{x}{a} \mathbb{1}_{\{x \leq \bar{a}\}} + \frac{\bar{x}}{a} \mathbb{1}_{\{x \geq \bar{a}\}}\right) da \\ &= x \int_x^1 \frac{d\bar{a}}{\bar{a}} + \bar{x} \int_{\bar{x}}^1 \frac{da}{a} \\ &= x \log(1/x) + \bar{x} \log(1/\bar{x}), \end{aligned}$$

and this expression extends continuously for $x \in [0, 1]$. In other words, h_ℓ converges point-wise to h_∞ .

However, since uniformly for all $x \in [0, 1]$ we have $h_\ell(x)^2 \leq 1$, the Dominated Convergence Theorem implies that also $h_\ell \rightarrow h_\infty$ in L^2 . □

With this preparation we will now determine the number of loose pairs asymptotically.

Lemma 2.7. *Let $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) = o(n)$.*

(a) *If both ℓ and ℓ' are constant, then we have*

$$\frac{B_{\ell, \ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\ell(t) h_{\ell'}(t) dt > 0.$$

(b) If $\ell = \ell(n) \rightarrow +\infty$, but ℓ' is a constant, then we have

$$\frac{B_{\ell, \ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\infty(t) h_{\ell'}(t) dt > 0.$$

(c) If $\ell' = \ell'(n) \rightarrow +\infty$, then we obtain

$$\frac{B_{\ell, \ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\infty(t)^2 dt = \frac{5}{6} - \frac{\pi^2}{18} = 0.2850\dots$$

Proof. Let $\Delta := \lfloor \frac{n-1}{\ell-1} \rfloor$ and $\Delta' := \lfloor \frac{n-1}{\ell'-1} \rfloor$. We enumerate the elements $(T, T') \in A_\ell \times A_{\ell'}$, with $T = (T(1), \dots, T(\ell))$ and $T' = (T'(1), \dots, T'(\ell'))$, by fixing the common differences $(\delta, \delta') \in [\Delta] \times [\Delta']$, and the unique intersection point $m \in [n]$ together with its positions $(\iota, \iota') \in [\ell] \times [\ell']$ within (T, T') . Then both ℓ -APs are to be contained in $[n]$ if and only if

$$1 \leq T(1) \wedge 1 \leq T'(1) \wedge T(\ell) \leq n \wedge T'(\ell') \leq n.$$

Expressing $T(1)$, $T'(1)$, $T(\ell)$, and $T'(\ell')$ in terms of m , ι , ι' , δ , and δ' , this is equivalent to

$$1 + \max\{(\iota - 1)\delta, (\iota' - 1)\delta'\} \leq m \leq n - \max\{(\ell - \iota)\delta, (\ell' - \iota')\delta'\}.$$

In other words, the number of valid choices for m is

$$(n - \max\{(\iota - 1)\delta, (\iota' - 1)\delta'\} - \max\{(\ell - \iota)\delta, (\ell' - \iota')\delta'\})_+,$$

with $x_+ := \max\{x, 0\} = x \mathbb{1}_{\{x \geq 0\}}$, and by summing over all choices for $(\iota, \iota', \delta, \delta') \in [\ell] \times [\ell'] \times [\Delta] \times [\Delta']$, we obtain

$$B_{\ell, \ell'} = \sum_{(\iota, \iota', \delta, \delta')} (n - \max\{(\iota - 1)\delta, (\iota' - 1)\delta'\} - \max\{(\ell - \iota)\delta, (\ell' - \iota')\delta'\})_+.$$

It turns out to be convenient to divide this quantity by n to obtain

$$\frac{B_{\ell, \ell'}}{n} = \sum_{(\iota, \iota', \delta, \delta')} f\left(\frac{\iota - 1}{\ell - 1}, \frac{\iota' - 1}{\ell' - 1}, \frac{(\ell - \iota)\delta}{n}, \frac{(\ell' - \iota')\delta'}{n}\right) \tag{2.6}$$

where the function $f: [0, 1]^4 \rightarrow [0, 1]$ is defined by

$$f(a, a', u, u') := (1 - \max\{au, a'u'\} - \max\{(1 - a)u, (1 - a')u'\})_+.$$

Now note that we have

$$\Delta = (1 \pm O(\ell/n)) \frac{n}{\ell - 1} \quad \text{and} \quad \Delta' = (1 \pm O(\ell'/n)) \frac{n}{\ell' - 1},$$

implying

$$\frac{(\ell - 1)\delta}{n} = (1 \pm O(\ell/n)) \frac{\delta}{\Delta} \quad \text{and} \quad \frac{(\ell' - 1)\delta'}{n} = (1 \pm O(\ell'/n)) \frac{\delta'}{\Delta'},$$

and thus it is not hard to show that there exists a constant $C > 0$ such that for all $1 \leq \iota \leq \ell$ and $1 \leq \iota' \leq \ell'$ we have

$$\left| f\left(\frac{\iota - 1}{\ell - 1}, \frac{\iota' - 1}{\ell' - 1}, \frac{(\ell - \iota)\delta}{n}, \frac{(\ell' - \iota')\delta'}{n}\right) - f\left(\frac{\iota - 1}{\ell - 1}, \frac{\iota' - 1}{\ell' - 1}, \frac{\delta}{\Delta}, \frac{\delta'}{\Delta'}\right) \right| \leq C \cdot \frac{\ell}{n}.$$

Furthermore, let

$$\nu_n(x, x') := \frac{1}{\Delta \Delta'} \sum_{(\delta, \delta')} \mathbb{1}_{\{x \leq \delta\}} \mathbb{1}_{\{x' \leq \delta'\}}$$

and observe that $\{d\nu_n\}_{n \in \mathbb{N}}$ converges weakly to the uniform measure on $[0, 1]^2$. Since f is bounded and continuous, we therefore have

$$\frac{B_{\ell, \ell'}}{n\Delta\Delta'} = (1 \pm o(1)) \sum_{(\iota, \iota')} \int_{[0, 1]^2} f\left(\frac{\iota - 1}{\ell - 1}, \frac{\iota' - 1}{\ell' - 1}, u, u'\right) dud u'. \tag{2.7}$$

The next goal is to deal with the positive part of the function f : we note that

$$(R - Q)_+ = R - \min\{R, Q\}$$

and so, for any $(a, a', u, u') \in [0, 1]^4$, by setting

$$\begin{aligned} R &:= \min\{1 - au, 1 - a'u'\}, \\ Q &:= \max\{(1 - a)u, (1 - a')u'\}, \end{aligned}$$

we obtain

$$\begin{aligned} f(a, a', u, u') &= \min\{1 - au, 1 - a'u'\} \\ &\quad - \min\{\min\{1 - au, 1 - a'u'\}, \max\{(1 - a)u, (1 - a')u'\}\}. \end{aligned}$$

Recall the integral representation

$$\min\{x, y\} = \int_0^{+\infty} \mathbb{1}_{\{t \leq x\}} \mathbb{1}_{\{t \leq y\}} dt,$$

which is valid for all $(x, y) \in \mathbb{R}_+^2$. We may express f as

$$\begin{aligned} f(a, a', u, u') &= \int_0^{+\infty} \mathbb{1}_{\{t \leq \min\{1 - au, 1 - a'u'\}\}} (1 - \mathbb{1}_{\{t \leq \max\{(1 - a)u, (1 - a')u'\}\}}) dt \\ &= \int_0^1 \mathbb{1}_{\{\max\{(1 - a)u, (1 - a')u'\} \leq t \leq \min\{1 - au, 1 - a'u'\}\}} dt \\ &= \int_0^1 \mathbb{1}_{\{(1 - a)u \leq t \leq 1 - au\}} \mathbb{1}_{\{(1 - a')u' \leq t \leq 1 - a'u'\}} dt \\ &= \int_0^1 \mathbb{1}_{\{u \leq \min\{t/\bar{a}, \bar{t}/a\}\}} \mathbb{1}_{\{u' \leq \min\{t/\bar{a}', \bar{t}'/a'\}\}} dt, \end{aligned}$$

using the convention that $1/0 = +\infty$, $\min\{x, +\infty\} = x$, and $\bar{x} := 1 - x$ for all $x \in [0, 1]$. Consequently, by integrating over $(u, u') \in [0, 1]^2$ and using Fubini's Theorem, we obtain

$$\int_{[0, 1]^2} f(a, a', u, u') dud u' = \int_0^1 \min\left\{\frac{t}{\bar{a}}, \frac{\bar{t}}{a}\right\} \min\left\{\frac{t}{\bar{a}'}, \frac{\bar{t}'}{a'}\right\} dt.$$

Hence, (2.7) simplifies to become

$$\frac{B_{\ell, \ell'}}{n\Delta\Delta'(\ell - 1)(\ell' - 1)} = (1 \pm o(1)) \int_{[0, 1]^3} \min\left\{\frac{t}{\bar{a}}, \frac{\bar{t}}{a}\right\} \min\left\{\frac{t}{\bar{a}'}, \frac{\bar{t}'}{a'}\right\} d\mu_\ell(a) d\mu_{\ell'}(a') dt,$$

where μ_ℓ and $\mu_{\ell'}$ are the measures defined in (2.1). Now, we observe that

$$n\Delta\Delta'(\ell - 1)(\ell' - 1) = (1 \pm o(1))n^3,$$

and so

$$\frac{B_{\ell, \ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\ell(t) h_{\ell'}(t) dt \stackrel{Cl. 2.5}{\geq} \int_{1/3}^{2/3} \frac{1}{4(\ell - 1)(\ell' - 1)} dt = \frac{1}{12(\ell - 1)(\ell' - 1)} > 0,$$

completing the proof of Lemma 2.7 when both ℓ and ℓ' are constant.

Assume now that ℓ' is a constant, but $\ell = \ell(n) \rightarrow +\infty$ with $\ell = o(\log n)$. Then by Claim 2.6 we have $h_\ell \rightarrow h_\infty$ in L^2 , furthermore, we have $\|h_{\ell'}\|_2 \leq 1$, hence

$$\left| \int_0^1 h_{\ell'}(t)(h_\infty(t) - h_\ell(t))dt \right| \leq \|h_{\ell'}\|_2 \cdot \|h_\infty - h_\ell\|_2 \rightarrow 0.$$

This implies that

$$\frac{B_{\ell,\ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\infty(t)h_{\ell'}(t)dt \stackrel{Cl.2.5}{\geq} \frac{1}{2(\ell' - 1)} \int_{1/3}^{2/3} h_\infty(t)dt \geq \frac{\log(3/2)}{3(\ell' - 1)} > 0,$$

completing the claim for this case.

Similarly, if $\ell = \ell(n) \rightarrow +\infty$ and $\ell' = \ell'(n) \rightarrow +\infty$ with $\ell' \leq \ell = o(\log n)$, then analogously to the previous case, we obtain

$$\frac{B_{\ell,\ell'}}{n^3} \xrightarrow{n \rightarrow +\infty} \int_0^1 h_\infty(t)^2 dt = \frac{5}{6} - \frac{\pi^2}{18} = 0.2850\dots,$$

where we evaluated the integral using SageMath [32]. □

Remark 2.8. The limits $\beta_{\ell,\ell'} := \lim_{n \rightarrow +\infty} B_{\ell,\ell'} n^{-3}$ can be computed explicitly based on their integral representation (and the help of SageMath) for specific choices of ℓ and ℓ' ; for instance, along the diagonal $\ell = \ell'$ we have

$$\beta_{3,3} = \frac{31}{48} \approx 0.6458; \quad \beta_{4,4} = \frac{130}{243} \approx 0.5350; \quad \beta_{5,5} = \frac{835}{1728} \approx 0.4832; \quad \dots$$

and similarly, we obtain $\beta_{4,3} = \frac{785}{1296} \approx 0.6057$, $\beta_{5,3} = \frac{335}{576} \approx 0.5816$, and also $\beta_{5,4} = \frac{1339}{2592} \approx 0.5166$. Further values are easily computed explicitly, however we do not believe that there exists a closed form expression for $\beta_{\ell,\ell'}$ in general.

2.2 Second moments

Given any subset $T \subseteq [n]$, we define

$$Z_T := \mathbb{1}_{\{T \subseteq [n]_p\}} - p^{|T|}$$

so $\mathbb{E}(Z_T) = 0$ for all $T \subset [n]$, and for any $3 \leq \ell = \ell(n) \leq n$ we set

$$\bar{X}_\ell := X_\ell - \mathbb{E}(X_\ell) = \sum_{T \in \mathcal{A}_\ell} Z_T.$$

First, we prove that the main contribution of the centred second moments comes from loose pairs, overlap pairs, or a combination of both.

Lemma 2.9. For $0 < p = p(n) = o(1)$ and any $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) = o(n)$ we have

$$\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'}) = (1 \pm o(1)) \left[B_{\ell,\ell'} p^{\ell+\ell'-1} + C_{\ell,\ell'} p^\ell \right].$$

In particular, since $C_{\ell,\ell} = A_\ell$, we have

$$\sigma_\ell = (1 \pm o(1)) \sqrt{B_\ell p^{2\ell-1} + A_\ell p^\ell}.$$

Proof. We observe that for any $r \in [\ell']$ and $(T, T') \in \mathcal{A}_\ell \times \mathcal{A}_{\ell'}$ with $|T \cap T'| = r$, we have

$$\mathbb{E}(Z_T Z_{T'}) = \mathbb{E}(\mathbb{1}_{\{T \cup T' \subseteq [n]_p\}} - p^{\ell+\ell'-r}) = p^{\ell+\ell'-r} - p^{\ell+\ell'} = (1 \pm o(1)) p^{\ell+\ell'-r},$$

while for any $(T, T') \in \mathcal{A}_\ell \times \mathcal{A}_{\ell'}$ with $|T \cap T'| = 0$ we have

$$\mathbb{E}(Z_T Z_{T'}) = \mathbb{E}(Z_T)\mathbb{E}(Z_{T'}) = 0.$$

By distinguishing the size of the intersection we obtain

$$\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'}) = (1 \pm o(1)) \sum_{r=1}^{\ell'} D_{\ell, \ell'}^{(r)} p^{\ell + \ell' - r},$$

and recall that by definition $D_{\ell, \ell'}^{(1)} = B_{\ell, \ell'}$ and $D_{\ell, \ell'}^{(\ell')} = C_{\ell, \ell'}$.

Therefore, we first consider the contribution of summands with $2 \leq r \leq \lfloor 2\ell'/3 \rfloor$. By the first estimate of Claim 2.4 we have

$$\sum_{r=2}^{\lfloor 2\ell'/3 \rfloor} D_{\ell, \ell'}^{(r)} p^{\ell + \ell' - r} = O(1) \cdot \sum_{r=2}^{\lfloor 2\ell'/3 \rfloor} n^2 \ell(\ell')^2 p^{\ell + \ell' - r} = O(n^2 \ell(\ell')^3 p^{\ell + \ell'/3}) = o(C_{\ell, \ell'} p^\ell),$$

where for the last estimate, we recall that $C_{\ell, \ell'} = \Theta(n^2(\ell - \ell' + 1)(\ell')^{-1})$ by Corollary 2.3, and observe that $\frac{\ell(\ell')^3 p^{\ell'/3}}{\ell - \ell' + 1} = o(1)$ for all constellations of ℓ and ℓ' , since $p = o(1)$.

Next, we consider the contribution of summands with $\lfloor 2\ell'/3 \rfloor + 1 \leq r \leq \ell' - 1$. By the second estimate of Claim 2.4 we obtain

$$\begin{aligned} \sum_{r=\lfloor 2\ell'/3 \rfloor + 1}^{\ell' - 1} D_{\ell, \ell'}^{(r)} p^{\ell + \ell' - r} &= O(1) \cdot \sum_{r=\lfloor 2\ell'/3 \rfloor + 1}^{\ell' - 1} n^2(\ell - r + 1)(\ell')^{-1}(\ell' - r + 1)p^{\ell + \ell' - r} \\ &= O(n^2(\ell - \ell' + 1)(\ell')^{-1}p^\ell) \cdot \sum_{i=0}^{\lfloor \ell'/3 \rfloor - 2} (i + 2)^2 p^{i+1} \\ &= o(C_{\ell, \ell'} p^\ell), \end{aligned}$$

since the last sum is of order $O(p) = o(1)$.

Hence, the main contribution to $\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'})$ comes from the summands for $r = 1$ and $r = \ell'$, i.e. we have

$$\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'}) = (1 \pm o(1)) \left[B_{\ell, \ell'} p^{\ell + \ell' - 1} + C_{\ell, \ell'} p^\ell \right],$$

as claimed by the first statement. As for the second statement, we recall that by definition $B_{\ell, \ell} = B_\ell$ and $C_{\ell, \ell} = C_\ell = A_\ell$. □

Before investigating the limiting correlation more in details, we remind the classical FKG inequality (see e.g. [20, thm. 2.12]). We say that a function $f : 2^{[n]} \rightarrow \mathbb{R}$ is increasing (resp. decreasing) if for all $A \subset B \subset [n]$, $f(A) \leq f(B)$ (resp. $f(A) \geq f(B)$).

Theorem 2.10 (Fortuin, Kasteleyn and Ginibre (1971)). *Consider a function of the random variables $\mathbb{1}_{\{i \in [n]_P\}}$ or more generally, suppose that one has a set $[n]_{p_1, \dots, p_n} \equiv [n]_P$ such that $\mathbb{P}(i \in [n]_P) = p_i$ independently of all other elements. Last, consider two random variables X_1, X_2 that are increasing or decreasing functions of $[n]_P$. Then,*

$$\mathbb{E}(X_1 X_2) \geq \mathbb{E}(X_1)\mathbb{E}(X_2) \tag{2.8}$$

In particular, if $\mathcal{Q}_1, \mathcal{Q}_2$ are two increasing (resp. decreasing) families of subsets of $[n]$, then

$$\mathbb{E}(\mathcal{Q}_1 \cap \mathcal{Q}_2 \subset [n]_P) \geq \mathbb{P}(\mathcal{Q}_1 \subset [n]_P)\mathbb{P}(\mathcal{Q}_2 \subset [n]_P)$$

An important application of this theorem concerns the random variable X defined by $X := \sum_{S \in \mathcal{S}} \mathbf{1}_{\{S \subset [n]_P\}}$ for a certain family \mathcal{S} of non-empty subsets of $[n]_P$. Note that every random variable $\mathbf{1}_{\{S \subset [n]_P\}}$ is increasing, hence, X is increasing.

Now, for any $3 \leq \ell' = \ell(n) < \ell = \ell(n) \leq n$ we define

$$\kappa_{\ell, \ell'} := \lim_{n \rightarrow +\infty} \frac{\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'})}{\sigma_\ell \sigma_{\ell'}} \tag{2.9}$$

and observe that $0 \leq \kappa_{\ell, \ell'} \leq 1$, by the FKG inequality and the Cauchy-Schwarz inequality. The following proof shows implicitly that $\kappa_{\ell, \ell'}$ is well-defined, i.e. the limit in (2.9) exists.

Lemma 2.11. *Let $0 < p = p(n) = o(1)$ and $3 \leq \ell' = \ell'(n) < \ell = \ell(n) = o(n)$.*

(a) *If $np^{\ell-1}\ell \rightarrow 0$, then*

$$\kappa_{\ell, \ell'} = 0;$$

(b) *if $np^{\ell-1}\ell \rightarrow c \in \mathbb{R}_+$, then*

$$0 < \kappa_{\ell, \ell'} < 1;$$

(c) *if $np^{\ell-1}\ell \rightarrow +\infty$ and ℓ' is a constant, then*

$$0 < \kappa_{\ell, \ell'} < 1;$$

(d) *if $np^{\ell-1}\ell \rightarrow +\infty$ and $\ell' = \ell'(n) \rightarrow +\infty$, then*

$$\kappa_{\ell, \ell'} = 1.$$

Proof. By Lemma 2.9, we have

$$\frac{[\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'})]^2}{\sigma_\ell^2 \sigma_{\ell'}^2} = (1 \pm o(1)) \frac{[B_{\ell, \ell'} p^{\ell+\ell'-1} + C_{\ell, \ell'} p^\ell]^2}{[B_\ell p^{2\ell-1} + A_\ell p^\ell] \cdot [B_{\ell'} p^{2\ell'-1} + A_{\ell'} p^{\ell'}]}.$$

First assume that $np^{\ell-1}\ell \rightarrow 0$, then we have

$$B_\ell p^{2\ell-1} = \Theta(n^3 p^{2\ell-1}) = o(n^2 \ell^{-1} p^\ell) = o(A_\ell p^\ell),$$

by Claim 2.1 and Lemma 2.7. Consequently, using $(a + b)^2 \leq 2a^2 + 2b^2 = O(a^2 + b^2)$, we see that

$$\frac{[\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'})]^2}{\sigma_\ell^2 \sigma_{\ell'}^2} = O(1) \cdot \left[\frac{(B_{\ell, \ell'} p^{\ell+\ell'-1})^2}{A_\ell p^\ell \cdot B_{\ell'} p^{2\ell'-1}} + \frac{(C_{\ell, \ell'} p^\ell)^2}{A_\ell p^\ell \cdot A_{\ell'} p^{\ell'}} \right].$$

Furthermore, using Claim 2.1 and Lemma 2.7 we obtain

$$\frac{(B_{\ell, \ell'} p^{\ell+\ell'-1})^2}{A_\ell p^\ell \cdot B_{\ell'} p^{2\ell'-1}} = O(1) \cdot np^{\ell-1}\ell = o(1),$$

and similarly, from Claim 2.1 and Corollary 2.3 we deduce

$$\frac{(C_{\ell, \ell'} p^\ell)^2}{A_\ell p^\ell \cdot A_{\ell'} p^{\ell'}} = O(1) \cdot \frac{p^{\ell-\ell'} \ell(\ell - \ell' + 1)^2}{\ell'} = o(1).$$

Hence, letting $n \rightarrow +\infty$ we obtain

$$\kappa_{\ell, \ell'} = \lim_{n \rightarrow +\infty} \frac{\mathbb{E}(\bar{X}_\ell \bar{X}_{\ell'})}{\sigma_\ell \sigma_{\ell'}} \leq 0,$$

as claimed since we already argued that $\kappa_{\ell, \ell'} \geq 0$ by the FKG inequality.

On the other hand, if $np^{\ell-1}\ell \rightarrow c \in \mathbb{R}_+$, then Claim 2.1, Corollary 2.3, and Lemma 2.7 imply

$$A_{\ell'}p^{\ell'} = \Theta(n^2(\ell')^{-1}p^{\ell'}) = o(n^3p^{2\ell'-1}) = o(B_{\ell'}p^{2\ell'-1}),$$

and

$$C_{\ell,\ell'}p^{\ell} = \Theta(n^2(\ell - \ell' + 1)(\ell')^{-1}p^{\ell}) = o(n^3p^{\ell+\ell'-1}) = o(B_{\ell,\ell'}p^{\ell+\ell'-1}).$$

Thus we obtain

$$\frac{[\mathbb{E}(\bar{X}_{\ell}\bar{X}_{\ell'})]^2}{\sigma_{\ell}^2\sigma_{\ell'}^2} = (1 \pm o(1)) \cdot \frac{(B_{\ell,\ell'}p^{\ell+\ell'-1})^2}{[B_{\ell}p^{2\ell-1} + A_{\ell}p^{\ell}] \cdot B_{\ell'}p^{2\ell'-1}}. \tag{2.10}$$

Now let $\varphi_{\ell} := h_{\ell}$ if ℓ is a constant, and $\varphi_{\ell} := h_{\infty}$ if $\ell = \ell(n) \rightarrow +\infty$; and define $\varphi_{\ell'}$ analogously. We note that both φ_{ℓ} and $\varphi_{\ell'}$ are L^2 -integrable. Next, we take the limit $n \rightarrow +\infty$ in (2.10) and note that Lemma 2.7 implies

$$\kappa_{\ell,\ell'}^2 = \lim_{n \rightarrow +\infty} \frac{[\mathbb{E}(\bar{X}_{\ell}\bar{X}_{\ell'})]^2}{\sigma_{\ell}^2\sigma_{\ell'}^2} = \frac{1}{1 + \gamma} \cdot \frac{\langle \varphi_{\ell}, \varphi_{\ell'} \rangle^2}{\|\varphi_{\ell}\|_2^2 \|\varphi_{\ell'}\|_2^2}$$

where

$$\gamma = \gamma(c, \ell) := \lim_{n \rightarrow +\infty} \frac{A_{\ell}p^{\ell}}{B_{\ell}p^{2\ell-1}} = \begin{cases} \frac{\ell}{2(\ell-1)c\|\varphi_{\ell}\|_2^2} & \text{if } \ell \text{ is finite} \\ \frac{1}{2c\|\varphi_{\infty}\|_2^2} & \text{otherwise} \end{cases}.$$

In particular, the Cauchy-Schwarz inequality implies

$$\kappa_{\ell,\ell'}^2 \leq \frac{1}{1 + \gamma} < 1,$$

since $\gamma > 0$. On the other hand, Lemma 2.7 also guarantees that $\langle \varphi_{\ell}, \varphi_{\ell'} \rangle > 0$ and this implies

$$\kappa_{\ell,\ell'} > 0,$$

completing the proof for the case $np^{\ell-1}\ell \rightarrow c \in \mathbb{R}_+$.

Assume now that $np^{\ell-1}\ell \rightarrow +\infty$, then $A_{\ell'}p^{\ell'} = o(B_{\ell'}p^{2\ell'-1})$, $A_{\ell}p^{\ell} = o(B_{\ell}p^{2\ell-1})$, and $C_{\ell,\ell'}p^{\ell} = \Theta(n^2(\ell - \ell' + 1)(\ell')^{-1}p^{\ell}) = o(n^3p^{\ell+\ell'-1}) = o(B_{\ell,\ell'}p^{\ell+\ell'-1})$, by Claim 2.1, Corollary 2.3, and Lemma 2.7. Therefore, we obtain

$$\kappa_{\ell,\ell'}^2 = \lim_{n \rightarrow +\infty} \frac{(B_{\ell,\ell'})^2}{B_{\ell} \cdot B_{\ell'}} = \frac{\langle \varphi_{\ell}, \varphi_{\ell'} \rangle^2}{\|\varphi_{\ell}\|_2^2 \|\varphi_{\ell'}\|_2^2}$$

from Lemmas 2.9 and 2.7, using the notation of φ_{ℓ} and $\varphi_{\ell'}$ as in the previous case. As before, we observe that $\langle \varphi_{\ell}, \varphi_{\ell'} \rangle > 0$ and this implies

$$\kappa_{\ell,\ell'} > 0.$$

It remains to distinguish two cases: first, if $\ell' = \ell'(n) \rightarrow +\infty$, then also $\ell = \ell(n) \rightarrow +\infty$ and thus $\varphi_{\ell} = \varphi_{\ell'} = h_{\infty}$, but then clearly $\langle h_{\infty}, h_{\infty} \rangle = \|h_{\infty}\|_2^2$, so $\kappa_{\ell,\ell'} = 1$.

Second, the case where ℓ' is a constant can be treated with the cases of equality of the Cauchy-Schwarz inequality. We recall that $\langle f, g \rangle = \|f\|\|g\|$ in a given \mathbb{R} -vector space E iff the two functions are linearly dependent in E , i.e. there exists $\lambda \in \mathbb{R}^*$ s.t. $f = \lambda g$ with $f, g \in E$. We then observe that $h_{\ell'}$ and h_{∞} are linearly independent in L^2 . To see this, let $\varepsilon = \varepsilon(\ell') > 0$ be a sufficiently small constant, and observe that $h_{\infty}(x)^2 \leq (2x + x \log(1/x))^2 \leq 9x$ for all $x \leq \varepsilon$ implying

$$\int_0^{\varepsilon} h_{\infty}(x)^2 dx \leq \frac{9}{2}\varepsilon^2;$$

however, by Lemma 2.5 (b), for sufficiently small $\varepsilon > 0$, we have $h_{\ell'}(x) \geq 1/(\ell' - 1)$ and thus

$$\int_0^\varepsilon h_{\ell'}(x)^2 dx \geq \frac{1}{(\ell' - 1)^2} \varepsilon.$$

Consequently, for any sufficiently small constant $\varepsilon > 0$ we obtain

$$\int_0^\varepsilon \left(\frac{h_{\ell'}(x)}{\|h_{\ell'}\|} \right)^2 dx \geq \frac{1}{(\ell' - 1)^2 \|h_{\ell'}\|^2} \varepsilon > \frac{9}{2 \|h_\infty\|^2} \varepsilon^2 \geq \int_0^\varepsilon \left(\frac{h_\infty(x)}{\|h_\infty\|} \right)^2 dx,$$

and so the functions $h_{\ell'}$ and h_∞ are not linearly dependent in L^2 , as claimed. Consequently, the Cauchy-Schwarz inequality is a strict inequality and we obtain

$$\kappa_{\ell, \ell'} = \frac{\langle \varphi_\ell, \varphi_{\ell'} \rangle^2}{\|\varphi_\ell\|_2^2 \|\varphi_{\ell'}\|_2^2} < 1,$$

completing the proof. □

3 Univariate fluctuations: proof of Theorem 1.1

In this section we focus on univariate fluctuations of A_ℓ , i.e. we prove the two statements of Theorem 1.1. First we treat the Poisson regime, where the result follows directly from an application of the Chen-Stein method and the preliminary computations performed in Section 2 (with $\ell' = \ell$). Likewise, the Gaussian approximation is a consequence of a classical normality criterion.

3.1 Poisson regime: proof of Theorem 1.1(a)

We start by introducing the notion of a dependency graph. We emphasize the fact that this definition is the one that fits our purpose, and that there can be many other such notions (see e.g. [14, 20]).

Definition 3.1. Let $(Y_i)_{1 \leq i \leq N}$ be a sequence of random variables (on a common probability space). A (simple) graph $G = (V, E)$ with vertex set $V = [N]$ is called a dependency graph for $(Y_i)_{i \in [N]}$ if and only if for all disjoint subsets $U, U' \subseteq V$ with $E(U, U') = \emptyset$ we have

$$(Y_i)_{i \in U} \text{ is independent of } (Y_i)_{i \in U'},$$

where $E(U, U') := \{(i, j) \in E : i \in U \text{ and } j \in U'\}$ denotes the set of edges between U and U' . We denote the neighbourhood of a vertex $i \in [N]$ by $\mathcal{N}(i) := \mathcal{N}_G(i) := \{j \in U : (i, j) \in E\}$ and let $\widehat{\mathcal{N}}(i) := \mathcal{N}(i) \cup \{i\}$.

The dependency graph relevant to this paper is the following: given $3 \leq \ell' = \ell'(n) \leq \ell = \ell(n) \leq n$ we consider the graph

$$\mathcal{G}_{\ell, \ell'} = \mathcal{G}_{\ell, \ell'}(n) := \left(\mathcal{A}_\ell \cup \mathcal{A}_{\ell'}, \left\{ (T, T') \in (\mathcal{A}_\ell \cup \mathcal{A}_{\ell'})^2 : |T \cap T'| \geq 1 \right\} \right). \quad (3.1)$$

In other words, the vertices represent APs and edges indicate that the corresponding APs intersect. Clearly, $\mathcal{G}_{\ell, \ell'}$ is a dependency graph of the family $(\mathbb{1}_{\{T \in [n]_p\}})_{T \in \mathcal{A}_\ell \cup \mathcal{A}_{\ell'}}$. Moreover, it is clearly not a bipartite graph. We remark that including the possibility $\ell = \ell'$ is to just cover the univariate case, in which $\mathcal{A}_\ell = \mathcal{A}_{\ell'}$ and so $\mathcal{A}_\ell \cup \mathcal{A}_{\ell'} = \mathcal{A}_\ell$.

We define the following two quantities associated with a dependency graph G of $(Y_i)_{1 \leq i \leq N}$:

$$\begin{aligned} \mathcal{V}_1(G) &:= \sum_{i=1}^N \sum_{j \in \tilde{\mathcal{N}}_G(i)} \mathbb{E}(Y_i)\mathbb{E}(Y_j), \\ \mathcal{V}_2(G) &:= \sum_{i=1}^N \sum_{j \in \mathcal{N}_G(i)} \mathbb{E}(Y_i Y_j). \end{aligned} \tag{3.2}$$

We use a variant of the Chen-Stein method due to Arratia, Goldstein, and Gordon [1] (in a slightly simplified form).

Theorem 3.2 (Theorem 1 in [1]). *Let $(Y_i)_{1 \leq i \leq N}$ be Bernoulli random variables of expectation $p_i := \mathbb{E}(Y_i) > 0$. Set*

$$S_N := \sum_{i=1}^N Y_i, \quad \text{and} \quad \zeta := \mathbb{E}(S_N) = \sum_{i=1}^N p_i.$$

Let G be a dependency graph of $(Y_i)_{1 \leq i \leq N}$, and $\mathcal{V}_1(G)$, $\mathcal{V}_2(G)$ as in (3.2). Let \mathcal{Y} be a Poisson random variable with mean $\mathbb{E}(\mathcal{Y}) := \zeta$. Then, for any $U \subset \mathbb{N}$,

$$|\mathbb{P}(S_N \in U) - \mathbb{P}(\mathcal{Y} \in U)| \leq \mathcal{V}_1(G) + \mathcal{V}_2(G).$$

Remark 3.3. The theorem given in [1] uses an additional quantity $\mathcal{V}_3(G)$ given by

$$\mathcal{V}_3(G) := \sum_{i=1}^N \mathbb{E} \left(\left| \mathbb{E} \left(Y_i - p_i \mid (Y_j)_{j \notin \tilde{\mathcal{N}}_G(i)} \right) \right| \right)$$

due to the use of a different notion of dependency graphs. This quantity is irrelevant for us, as we always have $\mathcal{V}_3(G) = 0$.

Proof of Theorem 1.1(a). We fix any $3 \leq \ell = \ell(n) \leq n$ and aim to apply Theorem 3.2 to the family $(\mathbb{1}_{\{T \in [n]_p\}})_{T \in \mathcal{A}_\ell}$. The corresponding dependency graph \mathcal{G}_ℓ was defined in (3.1). Clearly, for any $T \in \mathcal{A}_\ell$ we have $\mathbb{E}(\mathbb{1}_{\{T \subseteq [n]_p\}}) = p^\ell$ and thus

$$\mathcal{V}_1(\mathcal{G}_\ell) = \sum_{T \in \mathcal{A}_\ell} \sum_{T' \in \tilde{\mathcal{N}}_{\mathcal{G}_\ell}(T)} \mathbb{E}(\mathbb{1}_{\{T \subseteq [n]_p\}})\mathbb{E}(\mathbb{1}_{\{T' \subseteq [n]_p\}}) = p^{2\ell} \sum_{r=1}^{\ell} D_\ell^{(r)} = O(n^3 p^{2\ell}),$$

where the last equality holds due to Corollary 2.3, Claim 2.4, and Lemma 2.7.

Next, we note that $\mathbb{E}(\mathbb{1}_{\{T \subseteq [n]_p\}} \mathbb{1}_{\{T' \subseteq [n]_p\}}) = p^{2\ell-r}$ for all $1 \leq r \leq \ell - 1$ and $(T, T') \in \mathcal{D}_\ell^{(r)}$. Thus, we obtain

$$\begin{aligned} \mathcal{V}_2(\mathcal{G}_\ell) &= \sum_{T \in \mathcal{A}_\ell} \sum_{T' \in \mathcal{N}_G(T)} \mathbb{E}(\mathbb{1}_{\{T \subseteq [n]_p\}} \mathbb{1}_{\{T' \subseteq [n]_p\}}) = \sum_{r=1}^{\ell-1} D_\ell^{(r)} p^{2\ell-r} \\ &= O(1) \cdot [n^3 p^{2\ell-1} + n^2 \ell^3 p^{\ell+1}], \end{aligned}$$

where the last estimate holds due to Claim 2.4 and Lemma 2.7.

Combining these two bounds and using the assumption $n^2 p^\ell / (\ell - 1) \rightarrow c$ for some $c \in \mathbb{R}_+$ yields

$$\mathcal{V}_1(\mathcal{G}_\ell) + \mathcal{V}_2(\mathcal{G}_\ell) = O(n^{-1+2/\ell} + n^{-2/\ell}) = o(1).$$

The same bound holds when $\ell \rightarrow +\infty$, $\ell = o(\log n)$ and $p^{\ell^4} \rightarrow 0$. Thus Theorem 3.2 is applicable for the family $(\mathbb{1}_{\{T \subseteq [n]_p\}})_{T \in \mathcal{A}_\ell}$ and shows that for all $U \subseteq \mathbb{N}$ we have

$$\left| \mathbb{P} \left(\sum_{T \in \mathcal{A}_\ell} \mathbb{1}_{\{T \subseteq [n]_p\}} \in U \right) - \mathbb{P}(\text{Po}(\lambda) \in U) \right| \leq o(1),$$

where

$$\lambda := \lim_{n \rightarrow +\infty} \sum_{T \in \mathcal{A}_\ell} \mathbb{E}(\mathbb{1}_{\{T \subseteq [n]_p\}}) \stackrel{Cl. 2.1}{=} \lim_{n \rightarrow +\infty} (1 \pm o(1)) \frac{n^2 p^\ell}{2(\ell - 1)} = c/2$$

completing the proof of Theorem 1.1(a). □

Remark 3.4. If we do not suppose the assumption of Theorem 1.1(a), namely that $n^2 p^\ell \ell^{-1} = O(1)$, we still have a Poisson approximation with $\text{Po}(\lambda_n)$ where $\lambda_n := n^2 p^\ell \ell^{-1}$ provided that $\mathcal{V}_1(\mathcal{G}_\ell) + \mathcal{V}_2(\mathcal{G}_\ell) = o(1)$. This is the case if $n^3 p^{2\ell-1} \rightarrow 0$ and $n^2 p^{\ell+1} \ell^3 \rightarrow 0$, which is equivalent in the first case to $p \ll n^{-3/(2\ell-1)}$, and in the second case to $p \ll n^{-2/(\ell+1)} \ell^{-3/(\ell+1)}$. It is well known that a Poisson random variable with diverging parameter converges in distribution (after rescaling) to a Gaussian, hence, this case shows that we have a Gaussian regime for the range $n^{-2/\ell} \ell^{1/\ell} \ll p \ll \min\{n^{-3/(2\ell-1)}, n^{-2/(\ell+1)} \ell^{-3/(\ell+1)}\}$.

3.2 Gaussian regime: proof of Theorem 1.1(b)

For the normal approximation we apply a criterion due to Janson [19], which was then refined by Mikhailov [27]. This normality criterion is based on controlling mixed cumulants of sum of random variables by means of an associated dependency graph. We follow the notation of [20].

Theorem 3.5 (e.g. Theorem 6.21 in [20]). *Let $(X_{i,n})_{1 \leq i \leq N_n}$ be a family of random variables with dependency graph Γ_n (as defined in Definition 3.1) and suppose that there exist constants $\{C_r\}_{r \in \mathbb{N}}$ independent of n , and quantities M_n and Q_n such that*

$$\mathbb{E} \left(\sum_{i=1}^{N_n} |X_{i,n}| \right) \leq M_n, \tag{3.3}$$

and for all V of constant size (i.e. $|V|$ is independent of n), we have

$$\sum_{i \in \mathcal{N}(V)} \mathbb{E}(|X_{i,n}| | (X_{j,n})_{j \in V}) \leq C_{|V|} Q_n, \tag{3.4}$$

where $\mathcal{N}(V) := \cup_{i \in V} \mathcal{N}(i)$ as in Definition 3.1.

Let $S_n := \sum_{i=1}^{N_n} X_{i,n}$ and $\sigma_n^2 := \mathbb{V}(S_n)$. If there exists an $s > 2$ such that

$$\frac{M_n}{\sigma_n} \left(\frac{Q_n}{\sigma_n} \right)^{s-1} \xrightarrow{n \rightarrow +\infty} 0 \tag{3.5}$$

then, we have

$$\frac{S_n - \mathbb{E}(S_n)}{\sigma_n} \xrightarrow[n \rightarrow +\infty]{d} \text{N}(0, 1).$$

Note that the proof of Theorem 3.5 shows that the assumption (3.5) becomes weaker as s increases. However, we will see that for this application it is satisfied for any $s > 0$.

Proof of Theorem 1.1(b). In the setting of ℓ -APs we have $S_n := X_\ell$, and note that by Claim 2.1 we have $\mathbb{E}(X_\ell) = p^\ell A_\ell$, i.e. we may choose $M_n := p^\ell A_\ell$.

Now, for $V \subset \mathcal{A}_\ell$, let $\Lambda(V) := \cup_{T \in \mathcal{N}(V)} T \subset [n]$ be the set of points covered by APs in V . Recall that $\Xi_a = \mathbb{1}_{\{a \in [n]_p\}}$ are independent Bernoulli random variable with parameter p . We write $Z(V)$ for the LHS of (3.4) and observe that

$$\begin{aligned} Z(V) &= \sum_{T \in \mathcal{N}(V)} \mathbb{E} \left(\prod_{a \in T} \Xi_a \mid (\Xi_k)_{k \in \Lambda(V)} \right) = \sum_{T \in \mathcal{N}(V)} p^{\ell - |T \cap \Lambda(V)|} \prod_{a \in T \cap \Lambda(V)} \Xi_a \\ &\leq \sum_{T \in \mathcal{N}(V)} p^{\ell - |T \cap \Lambda(V)|} \end{aligned}$$

as Ξ_a takes values in $\{0, 1\}$. First, we consider APs $T \in \mathcal{N}(V)$ in “loose configurations”, i.e. $|T \cap \Lambda(V)| = 1$. Note that there are at most $O_{|V|}(n\ell)$ of these T and the contribution to $Z(V)$ of each of them is $p^{\ell-1}$. On the other hand, there are at most $O_{|V|}(\ell^4)$ APs $T \in \mathcal{N}(V)$ with $|T \cap \Lambda(V)| \geq 2$, and trivially, each of their contribution to $Z(V)$ is upper bounded by 1. Together this means that there exist constants $\{C_r\}_{r \in \mathbb{N}}$ and we may choose $Q_n := np^{\ell-1}\ell + \ell^4$ such that $Z(V) \leq C_{|V|}Q_n$ for all $V \subset \mathcal{A}_\ell$ of constant size.

Recall that Lemma 2.9 gives $\sigma_n = (1 \pm o(1))\sqrt{B_\ell p^{2\ell-1} + C_\ell p^\ell}$ with $C_\ell = A_\ell = \Theta(n^2\ell^{-1})$ and $B_\ell = \Theta(n^3)$ by Claim 2.1 and Lemma 2.7, respectively. Thus we have $\sigma_n = \Theta(\sqrt{n^2 p^\ell \ell^{-1} (1 + np^{\ell-1}\ell)})$ and we distinguish two cases:

If $np^{\ell-1}\ell \geq 10$, then we have $M_n/\sigma_n = O(n^{1/2}p^{1/2}\ell^{-1})$ and also $Q_n/\sigma_n \leq np^{\ell-1}\ell^5/\sigma_n = O(n^{-1/2}p^{-1/2}\ell^5)$. Thus, for any $s > 2$, we have

$$\frac{M_n}{\sigma_n} \left(\frac{Q_n}{\sigma_n} \right)^{s-1} = O \left((np)^{-(s-2)/2} \ell^{5(s-1)-1} \right) = o(1),$$

since $np \rightarrow +\infty$ polynomially in n and $\ell = o(\log n)$.

Otherwise, we have $np^{\ell-1}\ell \leq 10$ which implies $M_n/\sigma_n = O(np^{\ell/2}\ell^{-1/2})$ and $Q_n/\sigma_n = O(n^{-1}p^{-\ell/2}\ell^{9/2})$. Consequently, for any $s > 2$, we obtain

$$\frac{M_n}{\sigma_n} \left(\frac{Q_n}{\sigma_n} \right)^{s-1} = O \left((n^2 p^\ell)^{-(s-2)/2} \ell^{(9s-10)/2} \right).$$

Next, we recall that by Remark 3.4 we may additionally assume that p is not too small, e.g. $p \geq \varepsilon n^{-\max\{3/(2\ell-1), 2/(\ell+1)\}}$ for any $\varepsilon = \varepsilon(n) > 0$ with $\varepsilon \rightarrow 0$. It remains to observe that when ε is decreasing sufficiently slowly this implies that $n^2 p^\ell \gg e^{\Omega(\log n)/\ell}$. Since $\ell = o(\log n)$, it follows that (3.5) is satisfied and applying Theorem 3.5 completes the proof of Theorem 1.1(b). \square

4 Bivariate fluctuations: proof of Theorem 1.2

For the rest of this Section, it will be convenient to assume that $\ell_2 < \ell_1$. More precisely, we let $3 \leq \ell_2 = \ell_2(n) < \ell_1 = \ell_1(n)$ and $0 < p = p(n) < 1$ such that

$$p\ell_1^9 \xrightarrow{n \rightarrow +\infty} 0, \tag{4.1}$$

$$n^2 p^{\ell_1} \ell_1^{-9} \xrightarrow{n \rightarrow +\infty} +\infty, \tag{4.2}$$

$$\frac{\ell_1}{\log n} \xrightarrow{n \rightarrow +\infty} 0. \tag{4.3}$$

Our goal is to apply the method of moments (cf. Theorems 1.3 and 1.4), therefore we want to determine the asymptotics of the k -th moments $\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right]$ for all

$k \in \mathbb{N}$ and $u_{\ell_1}, u_{\ell_2} \in \mathbb{R}$. (We recall that $\sigma_{\ell_i} = \sqrt{\mathbb{E}(\bar{X}_{\ell_i}^2)}$ denotes the standard deviation of X_{ℓ_i} for $i \in \{1, 2\}$.) By definition we have

$$\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right] = \sum_{\mathbf{T} \in (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k} \left(\frac{u_{\ell_1}}{\sigma_{\ell_1}} \right)^{k_1(\mathbf{T})} \left(\frac{u_{\ell_2}}{\sigma_{\ell_2}} \right)^{k_2(\mathbf{T})} \mathbb{E} \left(\prod_{T \in \mathbf{T}} Z_T \right) \tag{4.4}$$

where $k_i(\mathbf{T}) := |\{T \in \mathbf{T} : |T| = \ell_i\}|$, for $i \in \{1, 2\}$, is the number of ℓ_i -APs in \mathbf{T} .

Remark 4.1. Note that despite our assumption that $\ell_1 \neq \ell_2$, our approach also includes the univariate scenario: for $3 \leq \ell = \ell(n) = o(\log n)$ and $0 < p = p(n) < 1$ such that $p\ell^9 \rightarrow 0$ and $n^2 p^\ell \ell^{-9} \rightarrow +\infty$, we obtain the k -th moment $\mathbb{E}(\bar{X}_\ell^k)$ by setting $\ell_2 = \ell$, $\ell_1 = 2\ell$, $u_{\ell_2} = 1$, and $u_{\ell_1} = 0$.³

Furthermore, we observe that in the univariate case the additional assumption (4.2) comes without loss of generality, since we already noticed in Remark 3.4 that $\bar{X}_\ell \sigma_\ell^{-1}$ has a Gaussian limit if $n^2 p^\ell \ell^{-1} \rightarrow +\infty$ but $n^2 p^\ell \ell^{-9} = O(1)$.

4.1 Main contribution to the moments

In (4.4) we expressed the k -th moment of an arbitrary linear combination of \bar{X}_{ℓ_1} and \bar{X}_{ℓ_2} as a sum ranging over k -tuples of APs, each of length ℓ_1 or ℓ_2 . We will now show that for even k the main contribution to this sum comes from k -tuples $\mathbf{T} = (T_1, \dots, T_k)$ with a certain matching structure, namely there exists a bijective self-inverse mapping $\nu : [k] \rightarrow [k]$ without fixed point (we will call such permutation a (*perfect*) *matching*) such that \mathbf{T} satisfies

$$\forall i \in [k]: T_i \cap T_{\nu(i)} \neq \emptyset \quad \wedge \quad T_i \cap \left(\bigcup_{j \in [k] \setminus \{i, \nu(i)\}} T_j \right) = \emptyset. \tag{4.5}$$

We write $\mathcal{F}_\nu(k)$ for the set of (ordered) k -tuples satisfying (4.5) for a given matching ν , and observe that any two distinct sets $\mathcal{F}_\nu(k)$ and $\mathcal{F}_{\nu'}(k)$, $\nu \neq \nu'$, are disjoint and can be mapped bijectively onto each other. Thus for any even k let ν^* be defined by

$$\nu^*(2i - 1) = 2i, \quad \forall i \in [k/2],$$

and note that there are precisely $(k - 1)!!$ many distinct matchings ν if k is even, and none at all if k is odd.

Let $F(k)$ denote the contribution of k -tuples in $\mathcal{F}(k) := \bigcup_\nu \mathcal{F}_\nu(k)$ to the k -th moment $\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right]$, and set $\mathcal{F}(k) := \emptyset$ for k odd. Then we let $\mathcal{G}(k) := (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k \setminus \mathcal{F}(k)$ for all $k \in \mathbb{N}$, and denote the contribution of $\mathcal{G}(k)$ by $G(k)$. In other words, we have

$$\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right] = F(k) + G(k), \tag{4.6}$$

where

$$F(k) := \sum_{\mathbf{T} \in \mathcal{F}(k)} \left(\frac{u_{\ell_1}}{\sigma_{\ell_1}} \right)^{k_1(\mathbf{T})} \left(\frac{u_{\ell_2}}{\sigma_{\ell_2}} \right)^{k_2(\mathbf{T})} \mathbb{E} \left(\prod_{T \in \mathbf{T}} Z_T \right)$$

$$G(k) := \sum_{\mathbf{T} \in \mathcal{G}(k)} \left(\frac{u_{\ell_1}}{\sigma_{\ell_1}} \right)^{k_1(\mathbf{T})} \left(\frac{u_{\ell_2}}{\sigma_{\ell_2}} \right)^{k_2(\mathbf{T})} \mathbb{E} \left(\prod_{T \in \mathbf{T}} Z_T \right).$$

³Note that the choice $\ell_1 = 2\ell$ is somewhat arbitrary, it simply ensures that $\ell_1 \geq \ell_2$ holds. Since we also set $u_{\ell_1} = 0$ this is in fact insubstantial.

We observe that by the previous argument we may express $F(k)$ for even k as

$$F(k) = (k - 1)!! \sum_{\mathbf{T} \in \mathcal{F}_{\nu^*}(k)} \prod_{i=1}^{k/2} \mathbb{E} \left(\frac{u_{|T_{2i-1}|} Z_{T_{2i-1}}}{\sigma_{|T_{2i-1}|}} \cdot \frac{u_{|T_{2i}|} Z_{T_{2i}}}{\sigma_{|T_{2i}|}} \right). \tag{4.7}$$

Note that the summation now ranges over ordered k -tuples of APs $\mathbf{T} \in \mathcal{F}_{\nu^*}(k)$, where the first AP is matched with the second, the third with the fourth, and in general any AP on an odd position is matched with its successor. Thus it is convenient to slightly change our perspective: from now on we will regard \mathbf{T} as an ordered $\frac{k}{2}$ -tuple of intersecting pairs of APs (T_{2i-1}, T_{2i}) . Each such intersecting pair falls into precisely one of three categories, either both T_{2i-1} and T_{2i} are ℓ_1 -APs, both are ℓ_2 -APs, or they form a mixed intersecting pair, i.e. one is an ℓ_1 -AP while the other is an ℓ_2 -AP. Formally, we define sets of labels by setting

$$\Theta_j(\mathbf{T}) := \{i \in [k/2] : |T_{2i-1}| = |T_{2i}| = \ell_j\}$$

for $j \in \{1, 2\}$ and

$$\Theta_3(\mathbf{T}) := [k/2] \setminus (\Theta_1(\mathbf{T}) \cup \Theta_2(\mathbf{T})).$$

Consequently, by parametrising according to Θ_1 and Θ_2 , the expression (4.7) turns into

$$F(k) = (k - 1)!! \sum_{\theta_1=0}^{k/2} \sum_{\theta_2=0}^{k/2-\theta_1} \binom{k/2}{\theta_1, \theta_2, \theta_3} u_{\ell_1}^{2\theta_1+\theta_3} u_{\ell_2}^{2\theta_2+\theta_3} \times \sum_{\substack{\mathbf{T} \in \mathcal{F}_{\nu^*}(k) \\ |\Theta_1(\mathbf{T})|=\theta_1 \\ |\Theta_2(\mathbf{T})|=\theta_2}} \prod_{i=1}^{k/2} \mathbb{E} \left(\frac{Z_{T_{2i-1}}}{\sigma_{|T_{2i-1}|}} \cdot \frac{Z_{T_{2i}}}{\sigma_{|T_{2i}|}} \right), \tag{4.8}$$

since there are precisely $\binom{k/2}{\theta_1, \theta_2, \theta_3}$ many possibilities for partitioning the set $[k/2]$ into sets of sizes θ_1, θ_2 and $\theta_3 := k/2 - \theta_1 - \theta_2$. Note that this means we have already fixed a partition of the labels $[k/2] = \bigcup_{j=1, \dots, 3} \Theta_j(\mathbf{T})$ when choosing \mathbf{T} in the last sum.

We observe that the first part of this expression is already very reminiscent of a multinomial formula. In the next lemma, we show that this intuition is well justified and demonstrate that the leading order term of $F(k)/(k - 1)!!$ is given by a multinomial with three summands (representing the three categories of intersections) and exponent $k/2$ (the length of \mathbf{T} as a tuple of intersecting pairs of APs).

Lemma 4.2. *Let $k \in 2\mathbb{N}$ and $u_{\ell_1}, u_{\ell_2} \in \mathbb{R}$, then we have*

$$F(k) = (1 \pm o(1))(k - 1)!! [u_{\ell_1}^2 + u_{\ell_2}^2 + 2u_{\ell_1}u_{\ell_2}\kappa_{\ell_1, \ell_2}]^{k/2}.$$

Proof. Let $k \in 2\mathbb{N}$, $u_{\ell_1}, u_{\ell_2} \in \mathbb{R}$, and fix an integer partition $k/2 = \theta_1 + \theta_2 + \theta_3$. Our first goal is to enumerate all the k -tuples $\mathbf{T} \in \mathcal{F}_{\nu^*}(k)$ satisfying $|\Theta_1(\mathbf{T})| = \theta_1$ and $|\Theta_2(\mathbf{T})| = \theta_2$. We are aiming to proceed pair-by-pair, i.e. for rounds $i = 1, \dots, k/2$ we enumerate all possible ways of embedding (T_{2i-1}, T_{2i}) into the set $[n]$. To do so, we need to be careful to avoid reusing points from $[n]$ in different rounds.

We now formalise this idea. Given an integer $m \in \mathbb{N}$ and $j \in \{1, 2\}$ we define the sets

$$\mathcal{M}(m, j) := \left\{ (T, T') \in \mathcal{A}_{\ell_j}^2 : |T \cap T'| = m \right\},$$

and similarly

$$\mathcal{M}(m, 3) := \left\{ (T, T') \in (\mathcal{A}_{\ell_1} \times \mathcal{A}_{\ell_2}) \cup (\mathcal{A}_{\ell_2} \times \mathcal{A}_{\ell_1}) : |T \cap T'| = m \right\}.$$

In other words, these are the sets of intersecting pairs of APs (of the three categories) which are embedded in $[n]$ and overlap in precisely m elements. Note that we have

$$\sum_m \sum_{(T,T') \in \mathcal{M}(m,j)} \frac{\mathbb{E}(Z_T Z_{T'})}{\sigma_{|T|} \sigma_{|T'|}} = \begin{cases} \mathbb{E}(\bar{X}_{\ell_1}^2) / \sigma_{\ell_1}^2 = 1, & \text{for } j = 1, \\ \mathbb{E}(\bar{X}_{\ell_2}^2) / \sigma_{\ell_2}^2 = 1, & \text{for } j = 2, \\ 2\mathbb{E}(\bar{X}_{\ell_1} \bar{X}_{\ell_2}) / (\sigma_{\ell_1} \sigma_{\ell_2}) \rightarrow 2\kappa_{\ell_1, \ell_2}, & \text{for } j = 3, \end{cases} \quad (4.9)$$

where the range of summation of the overlap parameter m is $\{1, \dots, \ell_1\}$ when $j = 1$ and $\{1, \dots, \ell_2\}$ when $j \in \{2, 3\}$. We observe that the factor 2 for $j = 3$ is due to symmetry as we consider *ordered* pairs (T, T') , see also the definition of $\mathcal{M}(m, 3)$.

Now in any round $i \in \{1, \dots, k/2\}$, we enumerate all possible choices for (T_{2i-1}, T_{2i}) by first choosing the size of their overlap, say m_i , and then selecting an embedded pair $(T, T') \in \mathcal{M}_i := \mathcal{M}(m_i, j_i)$, where j_i is given as the unique solution of $i \in \Theta_{j_i}(\mathbf{T})$.⁴ However, choosing $(T_{2i-1}, T_{2i}) = (T, T')$ may not be a valid choice, since $T \cup T'$ may contain elements from T_j for some $j \in [2i - 2]$ (thus violating (4.5) and the definition of $\mathcal{F}(k)$). Nonetheless, we claim that almost all of them are indeed valid. More formally, let

$$\mathcal{M}_i^* := \left\{ (T, T') \in \mathcal{M}_i : (T \cup T') \cap \left(\bigcup_{j=1}^{2i-2} T_j \right) = \emptyset \right\}$$

and note that $\left| \bigcup_{j=1}^{2i-2} T_j \right| \leq k\ell_1$ for all $i \in [k/2]$.

Claim 4.3. *For any $R \subseteq [n]$ of size at most $k\ell_1$ we have*

$$|\{(T, T') \in \mathcal{M}_i : (T \cup T') \cap R \neq \emptyset\}| = o(|\mathcal{M}_i|).$$

Before we prove Claim 4.3, we show how to complete the argument assuming this statement. Observe that we can now express (4.7) by selecting an overlap parameter m_i and an element from \mathcal{M}_i^* and thus obtain

$$F(k) = (k - 1)!! \sum_{\theta_1=0}^{k/2} \sum_{\theta_2=0}^{k/2-\theta_1} \binom{k/2}{\theta_1, \theta_2, \theta_3} u_{\ell_1}^{2\theta_1+\theta_3} u_{\ell_2}^{2\theta_2+\theta_3} \times \prod_{i=1}^{k/2} \left(\sum_{m_i} \sum_{(T,T') \in \mathcal{M}_i^*} \frac{\mathbb{E}(Z_T Z_{T'})}{\sigma_{|T|} \sigma_{|T'|}} \right). \quad (4.10)$$

Furthermore, as the contribution $\frac{\mathbb{E}(Z_T Z_{T'})}{\sigma_{|T|} \sigma_{|T'|}}$ from each term $(T, T') \in \mathcal{M}_i$ is the same, Claim 4.3 shows that the error introduced by replacing \mathcal{M}_i^* with \mathcal{M}_i in (4.10) is negligible: it is accounted for by a factor of $(1 \pm o(1))$. Consequently, by (4.9) we obtain

$$\begin{aligned} F(k) &= (1 \pm o(1))(k - 1)!! \sum_{\theta_1=0}^{k/2} \sum_{\theta_2=0}^{k/2-\theta_1} \binom{k/2}{\theta_1, \theta_2, \theta_3} (u_{\ell_1}^2)^{\theta_1} (u_{\ell_2}^2)^{\theta_2} (2u_{\ell_1} u_{\ell_2} \kappa_{\ell_1, \ell_2})^{\theta_3} \\ &= (1 \pm o(1))(k - 1)!! [u_{\ell_1}^2 + u_{\ell_2}^2 + 2u_{\ell_1} u_{\ell_2} \kappa_{\ell_1, \ell_2}]^{k/2}, \end{aligned}$$

as claimed.

Proof of Claim 4.3. Fix an arbitrary $R \subseteq [n]$ of size at most $k\ell_1$ and define the subsets of \mathcal{M}_i consisting of pairs of APs intersecting R by $\mathcal{M}'_i := \{(T, T') \in \mathcal{M}_i : (T \cup T') \cap R \neq \emptyset\}$. Note first that once T is fixed, the number of choices for T' with $|T \cap T'| \geq 2$ is at most

⁴Recall that the partition $[k/2] = \dot{\bigcup}_{j=1, \dots, 3} \Theta_j(\mathbf{T})$ is already fixed at this point.

$O(\ell_1^4)$, as T' is completely determined by choosing two elements in T (for which there are at most ℓ_1^2 choices) and deciding their positions within T' (also at most ℓ_1^2 choices).

We first deal with the case $m_i \geq 2$. We will see that $|\mathcal{M}_i| = \Omega(n^2/\ell_1)$ and $|\mathcal{M}'_i| = O(n\ell_1^5)$, and thus $|\mathcal{M}'_i| = o(|\mathcal{M}_i|)$ since $\ell_1 = o(\log n)$. Indeed, note that for every $2\ell_1$ -AP T'' , we can let $T := \{T''(1), \dots, T''(|T|)\}$ and $T' := \{T''(|T| - m_i + 1), \dots, T''(|T| + |T'| - m_i)\}$. Then $(T, T') \in \mathcal{M}_i$. Thus, $|\mathcal{M}_i| \geq A_{2\ell_1} = \frac{n^2}{2(2\ell_1 - 1)}(1 - o(1))$ by Claim 2.1. On the other hand, to obtain a pair (T, T') in \mathcal{M}'_i , we need to choose first some $x \in (T \cup T') \cap R$, which has at most $|R| \leq k\ell_1$ choices. Then the arithmetic progression containing x , say T , is determined by picking a common difference, for which there are at most n choices. Then by the observation above, the number of choices for T' with $|T \cap T'| \geq 2$ is $O(\ell_1^4)$. Therefore, $|\mathcal{M}'_i| = O(\ell_1 \cdot n \cdot \ell_1^4)$ as claimed.

We then deal with the case $m_i = 1$. Similarly, we show that $|\mathcal{M}_i| = \Omega(n^3\ell_1^{-2})$ and $|\mathcal{M}'_i| = O(n^2\ell_1)$, and hence $|\mathcal{M}'_i| = o(|\mathcal{M}_i|)$ since $\ell_1 = o(\log n)$. Indeed, to get a pair (T, T') in \mathcal{M}_i , we have at least A_{ℓ_1} choices to fix T and then, upon choosing some $x \in T$ as its intersection with T' , there are at least $\frac{n/2}{\ell_1 - 1}$ choices to choose the common difference of T' . This is because if $x \geq n/2$ ($x \leq n/2$ respectively), then we can find T' with x as the last (first respectively) element. Again there are at most $O(\ell_1^5)$ such T' intersecting with T at more than one place, we then have $|\mathcal{M}_i| \geq A_{\ell_1} \cdot \frac{n/2}{\ell_1 - 1} - O(\ell_1^5) = \Omega(n^3\ell_1^{-2})$, by Claim 2.1. On the other hand, a pair in \mathcal{M}'_i is determined by choosing their single intersection point with R and their common differences. So $|\mathcal{M}'_i| \leq |R| \cdot n \cdot n = O(n^2\ell_1)$, completing the proof of the claim. \square

As demonstrated earlier, this also completes the proof of Lemma 4.2. \square

4.2 Minor contribution to the moments

Next we turn our attention to k -tuples in $\mathcal{G}(k) = (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k \setminus \mathcal{F}(k)$, where $k \in \mathbb{N}$ and $\mathcal{F}(k) = \emptyset$ if k is odd.

Lemma 4.4. *Let $k \in \mathbb{N}$, we have*

$$G(k) = \sum_{\mathbf{T} \in \mathcal{G}(k)} \mathbb{E} \left(\prod_{i=1}^k \frac{u_{|T_i|} Z_{T_i}}{\sigma_{|T_i|}} \right) = o(1).$$

We start with some preparation. We will change the order of summation in an algorithmic fashion as described below. First we fix an arbitrary total order π of the set $\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2}$ such that all ℓ_1 -APs come before any ℓ_2 -AP, i.e. we have $\pi(T) < \pi(T')$ for all $T \in \mathcal{A}_{\ell_1}$ and $T' \in \mathcal{A}_{\ell_2}$. We now explore any (non-empty) finite collection of APs component-wise as follows. Roughly speaking, given \mathbf{T} , let H be an auxiliary k -vertex graph, in which each vertex represents an AP in \mathbf{T} and two vertices are adjacent if and only if the corresponding APs have non-empty intersection. Then we will explore $V(H)$, moving from one vertex to one of its neighbours according to the ordering π and start the search from a new component whenever the current one is exhausted. For $\mathbf{T} \in \bigcup_{k \in \mathbb{N}} (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k$, we set $|\mathbf{T}| := \inf \{k \geq 1 : \mathbf{T} \in (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k\}$. More precisely, we perform the following algorithm:

INPUT: $\mathbf{T} \in \bigcup_{k \in \mathbb{N}} (\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2})^k$.

- (I) Initialise the inactive list \mathcal{L}_i and active list \mathcal{L}_a : $\mathcal{L}_i \leftarrow \mathbf{T}$, $\mathcal{L}_a \leftarrow \emptyset$, and $j \leftarrow 1$.
- (II) Start a new component: If $\mathcal{L}_a = \emptyset$, then let $\mathcal{L}_a \leftarrow \{\min_{\pi} \mathcal{L}_i\}$.

(III) Set:

$$\begin{aligned}
 T_j &\leftarrow \min_{\pi} \mathcal{L}_a, \\
 s_j &\leftarrow |T_j|, \quad (\text{size of the current AP}) \\
 t_j &\leftarrow \left| T_j \cap \bigcup_{j'=1}^{j-1} T_{j'} \right|, \quad (\text{size of the overlap with previous APs}) \\
 \mathcal{C} &\leftarrow \{T \in \mathcal{L}_i : T \cap T_j \neq \emptyset\}. \quad (\text{current component})
 \end{aligned}$$

(IV) Update:

$$\begin{aligned}
 \mathcal{L}_a &\leftarrow (\mathcal{L}_a \cup \mathcal{C}) \setminus \{T_j\}, \\
 \mathcal{L}_i &\leftarrow \mathcal{L}_i \setminus \mathcal{C}.
 \end{aligned}$$

(V) If $j = |\mathbf{T}|$, then STOP; otherwise, set $j \leftarrow j + 1$ and return to step (II).

OUTPUT: $\pi(\mathbf{T}) := (T_1, \dots, T_{|\mathbf{T}|})$ and $\tau(\mathbf{T}) := (\mathbf{t}, \mathbf{s})$, where $\mathbf{t} = (t_1, t_2, \dots, t_{|\mathbf{T}|})$ and $\mathbf{s} = (s_1, s_2, \dots, s_{|\mathbf{T}|})$.

We remark that our algorithm resembles that of the, say, Depth/Breadth First Search algorithm on graphs. The difference is that in our algorithm within a connected component, we search APs according to the ordering π .

Example 4.5. We show the output of the algorithm on the case $n = 3$, $\ell_1 = 1$, $\ell_2 = 2$ and $k = 3$. The set of ℓ_1 -APs is given by $\{(1), (2), (3)\}$ and the set of ℓ_2 -APs by $\{(1, 2), (1, 3), (2, 3)\}$. The order π is given by $\{(1) < (2) < (3) < (1, 2) < (1, 3) < (2, 3)\}$; it obviously satisfies the condition that ℓ_1 -APs come before ℓ_2 -APs. We choose for instance $\mathbf{T} = ((2, 3), (1), (1, 3)) \in (\mathcal{A}_1 \cup \mathcal{A}_2)^3$. The successive inputs are given by

1. Iteration ($j = 1$)

$$\begin{aligned}
 \text{(I)} \quad &\mathcal{L}_a = \emptyset, \mathcal{L}_i = \mathbf{T}, \\
 \text{(II)} \quad &\mathcal{L}_a = \{(1)\}, \\
 \text{(III)} \quad &T_1 = (1), s_1 = 1, t_1 = 0, \mathcal{C}_1 = \{T \in \mathbf{T} : T \cap T_1 \neq \emptyset\} = \{(1), (1, 3)\}, \\
 \text{(IV)} \quad &\mathcal{L}_a = \{(1, 2)\}, \mathcal{L}_i = \{(2, 3)\}
 \end{aligned}$$

2. Iteration ($j = 2$)

$$\begin{aligned}
 \text{(III)} \quad &T_2 = (1, 3), s_2 = 2, t_2 = |\{1, 3\} \cap \{1\}| = 1, \mathcal{C}_2 = \{(2, 3)\}, \\
 \text{(IV)} \quad &\mathcal{L}_a = \{(2, 3)\}, \mathcal{L}_i = \emptyset,
 \end{aligned}$$

3. Iteration ($j = 3$)

$$\text{(III)} \quad T_3 = (2, 3), s_3 = 2, t_3 = |\{2, 3\} \cap \{1, 3\}| = 1, \text{ and } j = 3, \text{ hence we stop.}$$

Output: $\pi(\mathbf{T}) = \{T_1, T_2, T_3\} = \{(1), (1, 3), (2, 3)\}$, $\mathbf{t} = (0, 1, 1)$ and $\mathbf{s} = (1, 2, 2)$.

Note that any permutation \mathbf{T}' of the input \mathbf{T} will result in the same ordered tuple $\pi(\mathbf{T}') = (T_1, \dots, T_{|\mathbf{T}|})$. We now assume that $|\mathbf{T}| = k$. Observe that \mathbf{t} and \mathbf{s} satisfy

$$\forall i \in [k]: s_i \in \{\ell_1, \ell_2\}, \tag{4.11}$$

$$\forall i \in [k]: 0 \leq t_i \leq s_i, \tag{4.12}$$

$$\forall i \in [k]: \{t_i = 0\} \implies \{s_i = \ell_1\} \vee \{s_j = \ell_2, \forall j = i, \dots, k\}, \tag{4.13}$$

where (4.13) follows from the choice of π . An illustration of (4.13) is given in Example 4.5.

For $r \in \{0, 1, \dots, \ell_1\}$ and $j \in \{1, 2\}$ we define index sets

$$I_{r,\ell_j} := \{i \in [k] : t_i = r, s_i = \ell_j\}, \quad I_r := I_{r,\ell_1} \cup I_{r,\ell_2}.$$

Additionally, note that if the input \mathbf{T} is such that there exists $i \in [k-1]$ for which $t_i = t_{i+1} = 0$, then T_i is disjoint from $\bigcup_{j \in [k] \setminus \{i\}} T_j$ implying $\mathbb{E}\left(\prod_{j=1}^k Z_{T_j}\right) = \mathbb{E}Z_{T_i} \cdot \mathbb{E}\left(\prod_{j \in [k] \setminus \{i\}} Z_{T_j}\right) = 0$, i.e. such \mathbf{T} does not contribute to $G(k)$. Consequently, we have

$$\forall i \in [k-1] : t_i + t_{i+1} > 0. \tag{4.14}$$

Similarly, $t_k > 0$, since otherwise Z_{T_k} is independent from $(Z_{T_1}, \dots, Z_{T_{k-1}})$ and thus \mathbf{T} does not contribute to $G(k)$. We write

$$\mathcal{T}_k := \{\mathbf{t} \in \{0, 1, \dots, \ell_1\}^k : \mathbf{t} \text{ satisfies (4.14) and } t_k > 0\}$$

for the set of all type vectors of length k which do not contain two consecutive zeros and do not end in a zero. In particular, this implies that we may assume $|I_0| \leq \frac{k}{2} - 1$ for even k and $|I_0| \leq \frac{k-1}{2}$ for odd k , in other words, we have

$$|I_{0,\ell_1}| + |I_{0,\ell_2}| \leq \lceil k/2 \rceil - 1. \tag{4.15}$$

Next, for any type vector $\mathbf{t} \in \{0, 1, \dots, \ell_1\}^k$, we define the set of *valid size-type* vectors

$$\mathcal{S}_k(\mathbf{t}) := \{\mathbf{s} \in \{\ell_1, \ell_2\}^k : (\mathbf{t}, \mathbf{s}) \text{ satisfies (4.11), (4.12), and (4.13)}\}.$$

The main idea is to enumerate the sum in (4.4) by first choosing the vector $\mathbf{t} \in \{0, 1, \dots, \ell_1\}^k$, then a valid size-type vector $\mathbf{s} \in \mathcal{S}_k(\mathbf{t})$, and lastly a tuple $(T_1, \dots, T_k) \in \mathcal{G}(k)$ such that $\tau(T_1, \dots, T_k) = (\mathbf{t}, \mathbf{s})$. In terms of formula, we obtain

$$G(k) = \sum_{\mathbf{t} \in \mathcal{T}_k} \sum_{\mathbf{s} \in \mathcal{S}_k(\mathbf{t})} \sum_{\substack{\mathbf{T} \in \mathcal{G}(k) \\ \tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})}} \mathbb{E}\left(\prod_{T \in \mathbf{T}} \frac{u_{|T|} Z_T}{\sigma_{|T|}}\right) = \sum_{\mathbf{t} \in \mathcal{T}_k} \sum_{\mathbf{s} \in \mathcal{S}_k(\mathbf{t})} M_{\mathbf{t}, \mathbf{s}} \cdot \mu_{\mathbf{t}, \mathbf{s}}, \tag{4.16}$$

where

$$M_{\mathbf{t}, \mathbf{s}} := |\{\mathbf{T} \in \mathcal{G}(k) : \tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})\}|$$

denotes the number of tuples with given type vectors (\mathbf{t}, \mathbf{s}) , and

$$\mu_{\mathbf{t}, \mathbf{s}} := \left(\prod_{i \in [k]} \frac{u_{s_i}}{\sigma_{s_i}}\right) \frac{1}{M_{\mathbf{t}, \mathbf{s}}} \sum_{\substack{\mathbf{T} \in \mathcal{G}(k) \\ \tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})}} \mathbb{E}(Z_{T_1} \cdots Z_{T_k})$$

is the average contribution to $G(k)$ of a k -tuple with given type vectors (\mathbf{t}, \mathbf{s}) .

We first aim to bound the average contribution $\mu_{\mathbf{t}, \mathbf{s}}$.

Proposition 4.6. *Let $\mathbf{t} \in \mathcal{T}_k$ and $\mathbf{s} \in \mathcal{S}_k(\mathbf{t})$, then we have*

$$\mu_{\mathbf{t}, \mathbf{s}} = (1 \pm o(1)) \left(\prod_{i \in [k]} \frac{u_{s_i}}{\sigma_{s_i}}\right) p^{\sum_{i \in [k]} (s_i - t_i)}.$$

Proof. Let $\mathbf{T} = (T_1, \dots, T_k) \in \mathcal{G}(k)$ with $\tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})$. Here T_1, \dots, T_k are in the order corresponding to the output of the exploring algorithm, hence we have $|T_i| = s_i$. We see that

$$\mathbb{E}(Z_{T_1} \cdots Z_{T_k}) = \mathbb{E}\left(\prod_{i \in [k]} (\mathbb{1}_{\{T_i \subseteq [n]_p\}} - p^{|T_i|})\right) = \sum_{R \subseteq [k]} \left(\prod_{i \in R} -p^{s_i}\right) \mathbb{P}\left(\bigcup_{i \notin R} T_i \subseteq [n]_p\right).$$

First observe that the summand Q_\emptyset for $R = \emptyset$ is given by

$$Q_\emptyset := \mathbb{P} \left(\bigcup_{i \in [k]} T_i \subseteq [n]_p \right) = p^{\sum_{i=1}^k (s_i - t_i)},$$

as $s_i - t_i = |T_i \setminus \cup_{i' \leq i-1} T_{i'}|$ and so $\sum_{i \in [k]} (s_i - t_i) = |\cup_{i \in [k]} T_i|$. Thus, it only remains to show that the remaining (constantly many) summands are all of lower order.

Let $r \in [k]$ and fix an arbitrary subset $R \subseteq [k]$ of size r . The absolute value of its contribution to $\mathbb{E}(Z_{T_1} \cdots Z_{T_k})$ is equal to

$$Q_R := \left(\prod_{i \in R} p^{s_i} \right) \mathbb{P} \left(\bigcup_{i \notin R} T_i \subseteq [n]_p \right) = p^{\sum_{i \in R} s_i + |\cup_{i \notin R} T_i|}.$$

Note that

$$\sum_{i \in R} s_i + |\cup_{i \notin R} T_i| = \sum_{i \in R} |T_i| + |\cup_{i \notin R} T_i| \geq |\cup_{i \in [k]} T_i| = \sum_{i \in [k]} (s_i - t_i).$$

Furthermore, if this last inequality is not an equality, then

$$Q_R \leq p^{\sum_{i=1}^k (s_i - t_i) + 1} = o \left(p^{\sum_{i=1}^k (s_i - t_i)} \right),$$

i.e. Q_R is negligible compared to Q_\emptyset .

Next, suppose towards contradiction that the equality holds, so

$$\sum_{i \in R} |T_i| = |\cup_{i \in [k]} T_i| - |\cup_{i \notin R} T_i|.$$

But at the same time we have

$$\sum_{i \in R} |T_i| \geq |\cup_{i \in R} T_i| \geq |\cup_{i \in [k]} T_i| - |\cup_{i \notin R} T_i|,$$

and thus all intermediate inequalities above must be equalities. This happens for the first inequality when $\{T_i\}_{i \in R}$ are pairwise disjoint and for the second inequality when $(\cup_{i \in R} T_i) \cap (\cup_{i \notin R} T_i) = \emptyset$. But this in turn implies that for any $i \in R$, the set T_i is disjoint from $\cup_{j \neq i} T_j$, so $t_i = t_{i+1} = 0$, contradicting (4.14).

Because these bounds are uniform over the choice of the k -tuple \mathbf{T} the statement follows by taking the average. \square

We now aim at bounding the number of summands $M_{\mathbf{t}, \mathbf{s}}$. To do so, recall that in the dependency graph $\mathcal{G}_{\ell_1, \ell_2}$ defined in (3.1), each vertex represents an AP in $\mathcal{A}_{\ell_1} \cup \mathcal{A}_{\ell_2}$, and two vertices form an edge if and only if the corresponding APs have non-empty intersection.

Proposition 4.7. *For all $\mathbf{t} \in \mathcal{T}_k$ and $\mathbf{s} \in \mathcal{S}_k(\mathbf{t})$, we have*

$$M_{\mathbf{t}, \mathbf{s}} = O(1) \cdot \prod_{j=1}^{|\mathcal{I}_0|} \left[\binom{n^2}{s_{r_j}} \cdot (ns_{r_j+1})^{\mathbb{1}_{\{t_{r_j+1}=1\}}} (s_{r_j}^2 s_{r_j+1}^2)^{\mathbb{1}_{\{t_{r_j+1} \geq 2\}}} - \mathbb{1}_{\{t_{r_j+1}=s_{r_j+1}=s_{r_j}\}} \cdot \prod_{i=r_j+2}^{r_{j+1}-1} (n\ell_1)^{\mathbb{1}_{\{t_i=1\}}} (s_i^2 \ell_1^2)^{\mathbb{1}_{\{t_i \geq 2\}}} \right].$$

Proof. First, note that for any \mathbf{T} such that $\tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})$ and $\pi(\mathbf{T}) = (T_1, \dots, T_k)$, the component structure of the induced subgraph $\mathcal{G}_{\ell_1, \ell_2} \left[\bigcup_{i \in [k]} T_i \right]$ is already determined by the type-vector \mathbf{t} . More precisely, for $j = 1, \dots, |I_0|$, let

$$r_j := \min \{i \in [k] \setminus \{r_1, \dots, r_{j-1}\} : t_i = 0\}$$

denote the j -th zero entry of \mathbf{t} , and set $r_{|I_0|+1} := k + 1$. Note that $r_1 = 1$ and $\{T_{r_j}, \dots, T_{r_{j+1}-1}\}$ forms a component of $\mathcal{G}_{\ell_1, \ell_2} \left[\bigcup_{i \in [k]} T_i \right]$ for all $j = 1, \dots, |I_0|$.

We will construct tuples \mathbf{T} with $\tau(\mathbf{T}) = (\mathbf{t}, \mathbf{s})$ in the order given by its reordering $\pi(\mathbf{T}) = (T_1, \dots, T_k)$. In particular, this means that we consider one component of $\mathcal{G}_{\ell_1, \ell_2} \left[\bigcup_{i \in [k]} T_i \right]$ after the other. Let $j = 1, \dots, |I_0|$ and assume that T_1, \dots, T_{r_j-1} have already been chosen.

Observe that, by (4.14), the j -th component contains at least two APs T_{r_j} and T_{r_j+1} . As T_{r_j} starts a new component ($t_{r_j} = 0$), the number of choices for T_{r_j} is at most $A_{s_{r_j}} = O(n^2 s_{r_j}^{-1})$ by Claim 2.1. Next we choose T_{r_j+1} :

- (a) if $t_{r_j+1} = 1$, then the number of choices is at most $O(n s_{r_j})$, since there are at most s_{r_j} choices for the common vertex $x \in T_{r_j} \cap T_{r_j+1}$, at most s_{r_j+1} choices for the position of x within T_{r_j+1} and $O(n/s_{r_j+1})$ for the common difference of T_{r_j+1} ;
- (b) if $t_{r_j+1} = s_{r_j+1} = s_{r_j}$, then there is only one possibility $T_{r_j+1} = T_{r_j}$;
- (c) otherwise, T_{r_j+1} is determined by choosing two elements from T_{r_j} and their respective positions within T_{r_j+1} , which amounts to at most $O(s_{r_j}^2 s_{r_j+1}^2)$ many choices.

Similarly, for any remaining $i = r_j + 2, \dots, r_{j+1} - 1$ (there might be none), we use the following bounds on the number of choices for T_i :

- (a) if $t_{r_j+1} = 1$, then the number of choices is at most $O(n \ell_1)$, since there are at most $O(\ell_1)$ choices for the common vertex $x \in T_i \cap (T_{r_j} \cup \dots \cup T_{i-1})$, at most s_i choices for the position of x within T_i and $O(n/s_i)$ for the common difference of T_i ;
- (b) otherwise, T_i is determined by choosing two elements from $T_{r_j} \cup \dots \cup T_{i-1}$ and their respective positions within T_i , which amounts to at most $O(\ell_1^2 s_i^2)$ many choices.

The claim follows by multiplying for all $j = 1, \dots, |I_0|$ and $i = r_j, \dots, r_{j+1} - 1$. □

With this preparation we are now ready to prove Lemma 4.4. We will bound the contribution of each k -tuple to $G(k) = \sum_{\mathbf{t}, \mathbf{s}} \mu_{\mathbf{t}, \mathbf{s}} M_{\mathbf{t}, \mathbf{s}}$ from above component-wise

Proof of Lemma 4.4. First observe that Lemma 2.9 implies that for any $\ell \geq 3$ we have

$$\sigma_\ell^{-1} = O(A_\ell p^\ell)^{-1/2} \stackrel{C.2.1}{=} O(n^{-1} p^{-\ell/2} \ell^{1/2}), \tag{4.17}$$

and also

$$\sigma_\ell^{-1} = O(B_\ell p^{2\ell-1})^{-1/2} \stackrel{L.2.7}{=} O(n^{-3/2} p^{-\ell+1/2}). \tag{4.18}$$

Using Propositions 4.6 and 4.7 the expression in (4.16) becomes

$$G(k) = O(1) \cdot \sum_{\mathbf{t} \in \mathcal{T}_k} \sum_{\mathbf{s} \in \mathcal{S}_k(\mathbf{t})} \prod_{i \in [k]} g_{\mathbf{t}, \mathbf{s}}(i) \sigma_{s_i}^{-1}, \tag{4.19}$$

where

$$g_{\mathbf{t},\mathbf{s}}(i) := \begin{cases} n^2 s_i^{-1} p^{s_i} & \text{if } t_i = 0; \\ n s_i p^{s_i - 1} & \text{if } t_i = 1 \wedge t_{i-1} = 0; \\ n \ell_1 p^{s_i - 1} & \text{if } t_i = 1 \wedge t_{i-1} > 0; \\ 1 & \text{if } t_i = s_i = s_{i-1} \wedge t_{i-1} = 0; \\ s_i^2 s_{i-1}^2 p^{s_i - t_i} & \text{if } 2 \leq t_i \leq s_i - 1 \wedge t_{i-1} = 0; \\ s_i^2 s_{i-1}^2 & \text{if } t_i = s_i \wedge s_i \neq s_{i-1} \wedge t_{i-1} = 0; \\ s_i^2 \ell_1^2 p^{s_i - t_i} & \text{if } t_i \geq 2 \wedge t_{i-1} > 0. \end{cases}$$

Moreover, we recall the notation $r_j = \min \{i \in [k] \setminus \{r_1, \dots, r_{j-1}\} : t_i = 0\}$ and $r_{|I_0|+1} := k + 1$ used in the proof of Proposition 4.7. These indices split the interval $[k]$ into $|I_0|$ parts, i.e. $[k] = \bigcup_{j=1}^{|I_0|} \{r_j, \dots, r_{j+1} - 1\}$, where each part has size at least two as, by (4.14), \mathbf{t} does not have consecutive zeros. Now, fix any $j \in [|I_0|]$. We first bound the first two factors together:

(a) If $t_{r_j+1} = 1$, then we have

$$\frac{g_{\mathbf{t},\mathbf{s}}(r_j) g_{\mathbf{t},\mathbf{s}}(r_j + 1)}{\sigma_{s_{r_j}} \sigma_{s_{r_j+1}}} = n^2 s_{r_j}^{-1} p^{s_{r_j}} \cdot n s_{r_j+1} p^{s_{r_j+1} - 1} \cdot \sigma_{s_{r_j}}^{-1} \sigma_{s_{r_j+1}}^{-1} \stackrel{(4.18)}{=} O(1),$$

because the validity condition (4.13) implies $s_{r_j+1} \leq s_{r_j}$, since $t_{r_j} = 0$.

(b) If $2 \leq t_{r_j+1} \leq s_{r_j+1} - 1$, then we have

$$\begin{aligned} \frac{g_{\mathbf{t},\mathbf{s}}(r_j) g_{\mathbf{t},\mathbf{s}}(r_j + 1)}{\sigma_{s_{r_j}} \sigma_{s_{r_j+1}}} &= n^2 s_{r_j}^{-1} p^{s_{r_j}} \cdot s_{r_j}^2 s_{r_j+1}^2 p^{s_{r_j+1} - t_{r_j+1}} \cdot \sigma_{s_{r_j}}^{-1} \sigma_{s_{r_j+1}}^{-1} \\ &\stackrel{(4.17)}{=} O\left(p^{1+(s_{r_j} - s_{r_j+1})/2} \ell_1^4\right) = o(\ell_1^{-1}) \end{aligned}$$

since $p \ell_1^5 \rightarrow 0$, by assumption (4.1).

(c) If $t_{r_j+1} = s_{r_j+1} \leq s_{r_j} - 1$, then we have

$$\begin{aligned} \frac{g_{\mathbf{t},\mathbf{s}}(r_j) g_{\mathbf{t},\mathbf{s}}(r_j + 1)}{\sigma_{s_{r_j}} \sigma_{s_{r_j+1}}} &= n^2 s_{r_j}^{-1} p^{s_{r_j}} \cdot s_{r_j}^2 s_{r_j+1}^2 \cdot \sigma_{s_{r_j}}^{-1} \sigma_{s_{r_j+1}}^{-1} \\ &\stackrel{(4.17)}{=} O\left(p^{(s_{r_j} - s_{r_j+1})/2} \ell_1^4\right) = o(1), \end{aligned}$$

since $p^{1/2} \ell_1^4 \rightarrow 0$, by assumption (4.1).

(d) If $t_{r_j+1} = s_{r_j+1} = s_{r_j}$, then we have

$$\frac{g_{\mathbf{t},\mathbf{s}}(r_j) g_{\mathbf{t},\mathbf{s}}(r_j + 1)}{\sigma_{s_{r_j}} \sigma_{s_{r_j+1}}} = n^2 s_{r_j}^{-1} p^{s_{r_j}} \cdot \sigma_{s_{r_j}}^{-1} \sigma_{s_{r_j+1}}^{-1} \stackrel{(4.17)}{=} O(1).$$

We now treat any (potentially) remaining indices $i = r_j + 2, \dots, r_{j+1} - 1$ and estimate $g_{\mathbf{t},\mathbf{s}}(i) \sigma_{s_i}^{-1}$ one by one.

(a) If $t_i = 1$, then we have

$$g_{\mathbf{t},\mathbf{s}}(i) \sigma_{s_i}^{-1} = n \ell_1 p^{s_i - 1} \sigma_{s_i}^{-1} \stackrel{(4.18)}{=} O((np)^{-1/2} \ell_1) = o(1),$$

since $3 \leq \ell_1 = o(\log n)$ and $np = \Omega(n^{1-2/\ell_1}) = n^{\Omega(1)}$, by assumption (4.2).

(b) If $2 \leq t_i \leq s_i - 1$, then we have

$$g_{\mathbf{t},\mathbf{s}}(i)\sigma_{s_i}^{-1} \leq s_i^2 \ell_1^2 p^{s_i-t_i} \sigma_{s_i}^{-1} = O(p\ell_1^4) = o(\ell_1^{-1}),$$

since $\sigma_{s_i}^{-1} = O(1)$ by assumption (4.2) and because $p\ell_1^5 \rightarrow 0$, by assumption (4.1).

(c) However, if $t_i = s_i$, then we have

$$g_{\mathbf{t},\mathbf{s}}(i)\sigma_{s_i}^{-1} \leq s_i^2 \ell_1^2 \sigma_{s_i}^{-1} = O(\sigma_{\ell_1}^{-1} \ell_1^4) \stackrel{(4.17)}{=} O(n^{-1} p^{-\ell_1/2} \ell_1^{9/2}) = o(1),$$

since $n^2 p^{\ell_1} \ell_1^{-9} \rightarrow +\infty$, by assumption (4.2).

Next, we observe that by (4.15), we have $|I_0| \leq \lceil k/2 \rceil - 1$, implying that there must be at least one $j \in |I_0|$ such that there exists an integer i_0 satisfying $r_j + 2 \leq i_0 \leq r_{j+1} - 1$. But then, the previous computation shows that the corresponding factor $g_{\mathbf{t},\mathbf{s}}(i_0)\sigma_{s_{i_0}}^{-1}$ must be small. More precisely, we have

$$g_{\mathbf{t},\mathbf{s}}(i_0)\sigma_{s_{i_0}}^{-1} = o(1) \cdot \ell_1^{-\mathbb{1}_{\{2 \leq t_{i_0} \leq s_{i_0} - 1\}}}.$$

Consequently, from (4.19) and multiplying the bounds for all $i \in [k]$ we obtain

$$G(k) = O(1) \cdot \sum_{\mathbf{t} \in \mathcal{T}_k} \sum_{\mathbf{s} \in \mathcal{S}_k(\mathbf{t})} o(\ell_1^{-Q(\mathbf{t},\mathbf{s})}),$$

where

$$Q(\mathbf{t}, \mathbf{s}) := |\{i \in [k] : 2 \leq t_i \leq s_i - 1\}|.$$

Last, for any $q \in \{0, 1, \dots, k\}$, the number of summands with $Q(\cdot) = q$ is at most $\ell_1^q 3^{k-q} 2^k = O(\ell_1^q)$ yielding

$$G(k) = o(1),$$

thereby completing the proof of Lemma 4.4. □

4.3 Completing the argument: application of the method of moments

It remains to apply the method of moments to show the convergence to a (bivariate) Gaussian distribution.

Proof of Theorem 1.2. Recall from (4.6) that we have

$$\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right] = F(k) + G(k),$$

for all $k \in \mathbb{N}$ and $u_{\ell_1}, u_{\ell_2} \in \mathbb{R}$. Furthermore, we have computed the asymptotics for $F(k)$ and $G(k)$ in Lemmas 4.2 and 4.4, implying that for even k we have

$$\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right] = (1 \pm o(1))(k-1)!! [u_{\ell_1}^2 + u_{\ell_2}^2 + 2\kappa_{\ell_1, \ell_2} u_{\ell_1} u_{\ell_2}]^{k/2},$$

and for odd k we have

$$\mathbb{E} \left[\left(u_{\ell_1} \frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}} + u_{\ell_2} \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right)^k \right] = o(1).$$

Letting $n \rightarrow +\infty$ we obtain the k -th moments of the bivariate standard Gaussian distribution with covariance κ_{ℓ_1, ℓ_2} . Hence, Theorems 1.3 and 1.4 imply that

$$\left(\frac{\bar{X}_{\ell_1}}{\sigma_{\ell_1}}, \frac{\bar{X}_{\ell_2}}{\sigma_{\ell_2}} \right) \xrightarrow[n \rightarrow +\infty]{d} \mathbb{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \kappa_{\ell_1, \ell_2} \\ \kappa_{\ell_1, \ell_2} & 1 \end{pmatrix} \right).$$

The distinction of the different regimes in Theorem 1.2 follows from Lemma 2.11, completing the proof. \square

The same proof also applies for the study of univariate fluctuations.⁵

Alternative proof of Theorem 1.1(b). For $3 \leq \ell = \ell(n) = o(\log n)$ and $0 < p = p(n) < 1$ such that $p\ell^9 \rightarrow 0$ and $n^2p\ell^{-9} \rightarrow +\infty$, we obtain

$$\mathbb{E} \left[\left(\frac{\bar{X}_\ell}{\sigma_\ell} \right)^k \right] = \begin{cases} (1 \pm o(1))(k-1)!! & \text{for } k \text{ even,} \\ o(1) & \text{if for } k \text{ odd,} \end{cases}$$

from Lemmas 4.2 and 4.4 by setting $\ell_2 = \ell$, $\ell_1 = 2\ell$, $u_{\ell_2} = 1$, and $u_{\ell_1} = 0$. Letting $n \rightarrow +\infty$ Theorem 1.3 shows that

$$\frac{\bar{X}_\ell}{\sigma_\ell} \xrightarrow[n \rightarrow +\infty]{d} N(0, 1),$$

as claimed. \square

5 Concluding remarks

The main topic at stake in this article was to study the joint distribution of the numbers of APs of different length in some random subsets M of the integers. In the most general setup, we would like to understand the growth behaviour of the family $\{X_\ell\}_{3 \leq \ell \leq n}$ where $X_\ell = X_\ell(M)$ denotes the number of ℓ -APs of integers which are (entirely) contained in M . Here, we took a first step in this direction by determining the joint limiting distribution of (X_{ℓ_1}, X_{ℓ_2}) in $M = [n]_p$ for a significant range of parameters p and $3 \leq \ell_2 < \ell_1 = o(\log n)$. We believe that our approach should also allow us to determine the limiting distribution of r -tuples $(X_{\ell_1}, X_{\ell_2}, \dots, X_{\ell_r})$ for $r \geq 3$ (within the intersection of their respective Gaussian regimes), hence, to give a functional Central Limit Theorem for e.g. $(X_{\lfloor s\ell \rfloor})_{s \in [0,1]}$ with $\ell = \ell(n) = o(\log n)$. In particular, it would be interesting to know whether for some constants $\ell_1, \ell_2, \dots, \ell_r$, with (constant) $r \geq 3$, the Gaussian limit becomes degenerate. We observed it for $r = 2$ when $\ell_1 = \ell_1(n), \ell_2 = \ell_2(n) \rightarrow +\infty$ sufficiently slowly: X_{ℓ_1} and X_{ℓ_2} are then either asymptotically uncorrelated or converge to the same Gaussian random variable (after re-normalisation).

Furthermore, recall that Theorem 1.2 uses the assumption $n^2p\ell_1\ell_1^{-9} \rightarrow +\infty$ which guarantees that both X_{ℓ_1} and X_{ℓ_2} are within their respective Gaussian regimes. One may thus ask what happens for smaller values of p . At least heuristically, our results for the overlap pair regime (i.e. $np^{\ell_1-1}\ell_1 \rightarrow 0$) suggest that a good candidate for the joint limit consists of two independent random variables having the appropriate marginal distributions (Gaussian or Poisson) determined in Theorem 1.1.

Throughout the article, we focused on ℓ -APs where $\ell = o(\log n)$, the reason being that typically the random set $[n]_p$ will not contain any longer APs as long as $p = o(1)$. In order to witness any ℓ -APs with $\ell/\log n \rightarrow +\infty$ we would need to consider $p = p(n) \rightarrow 1$. Borrowing some intuition from Gao and Sato's work [17] on large matchings in the random graph $G(n, p)$ – namely the log-normal paradigm of Gao [16] – we might expect to see another change of regime to a Log-normal limiting distribution for very long APs. However, in this regime, various estimates derived in this paper cease to hold and we leave this as an open problem.

Another question of interest concerns the behaviour of the joint cumulants of (X_{ℓ_1}, X_{ℓ_2}) in the various regimes encountered here. In the Gaussian regime, since the moments of the rescaled random variables converge to the Gaussian moments, their

⁵Albeit with the mild additional assumption $p\ell^9 \rightarrow 0$ for technical reasons.

cumulants of order $r \geq 3$ converge to 0. One can ask if the BFS-type coding allows to see such a behaviour in a fine way, for instance with an asymptotic expansion.

Lastly, we would like to move in a slightly different direction: let $0 < s < t$ and consider the coupling $[[tn]]_p = [[sn]]_p \cup \{[sn] + 1, \dots, [tn]\}_p$ for any $p \in [0, 1]$. What can be said about the joint distribution of $(X_\ell([[sn]]_p), X_\ell([[tn]]_p))$? More generally, does the random process $(X_\ell([[tn]]_p))_{t \geq 0}$ satisfy a functional central limit theorem? What about $(X_{[s\ell]}([[tn]]_p))_{s, t \geq 0}$ for $\ell = \ell(n) = o(\log n)$?

References

- [1] Richard Arratia, Larry Goldstein, and Louis Gordon, *Two moments suffice for Poisson approximations: the Chen-Stein method*, Ann. Probab. **17** (1989), no. 1, 9–25. MR-0972770
- [2] József Balogh, Robert Morris, Wojciech Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), 669–709. MR-3327533
- [3] Bhaswar B. Bhattacharya, Shirshendu Ganguly, Xuancheng Shao, and Yufei Zhao, *Upper tails for arithmetic progressions in a random set*, arXiv:1605.02994 (2016).
- [4] Patrick Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1995, A Wiley-Interscience Publication. MR-1324786
- [5] Béla Bollobás, *To prove and conjecture: Paul Erdős and his mathematics*, Amer. Math. Monthly **105** (1998), no. 3, 209–237. MR-1615568
- [6] Béla Bollobás, Oliver Cooley, Mihyun Kang, and Christoph Koch, *Jigsaw percolation on random hypergraphs*, J. Appl. Probab. **54** (2017), no. 4, 1261–1277. MR-3731294
- [7] Jop Briët and Sivakanth Gopi, *Gaussian width bounds with applications to arithmetic progressions in random settings*, arXiv:1711.05624 (2017).
- [8] Sourav Chatterjee, *An introduction to large deviations for random graphs*, Bull. Amer. Math. Soc. (N.S.) **53** (2016), no. 4, 617–642. MR-3544262
- [9] Sourav Chatterjee and Amir Dembo, *Nonlinear large deviations*, Adv. Math. **299** (2016), 396–450. MR-3519474
- [10] David Conlon, W. Timothy Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. **184** (2016), no. 2, 367–454. MR-3548529
- [11] Harald Cramér, *On the order of magnitude of the difference between consecutive prime numbers*, Acta Arith. **2** (1936), no. 1, 23–46.
- [12] J. P. G. Lejeune Dirichlet, *Beweis des satzes, dass jede unbegrenzte arithmetische progression, deren erstes glied und differenz ganze zahlen ohne gemeinschaftlichen factor sind, unendlich viele primzahlen enthält*, Abhandlungen der Königlich Preussischen Akademie der Wissenschaften **45** (1837), 81.
- [13] Ronen Eldan, *Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations*, Geom. Funct. Anal. **28** (2018), no. 6, 1548–1596. MR-3881829
- [14] Valentin Féray, *Weighted dependency graphs*, Electron. J. Probab. **23** (2018), Paper No. 93, 65. MR-3858921
- [15] Zoltán Füredi, János Komlós, *The eigenvalues of random symmetric matrices*, Combinatorica **1** (1981), no. 3, 233–241. MR-0637828
- [16] Pu Gao, *Distributions of sparse spanning subgraphs in random graphs*, SIAM J. Discrete Math. **27** (2013), no. 1, 386–401. MR-3032925
- [17] Pu Gao and Cristiane M. Sato, *A transition of limiting distributions of large matchings in random graphs*, J. Combin. Theory Ser. B **116** (2016), 57–86. MR-3425237
- [18] Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. **167** (2008), no. 2, 481–547. MR-2415379
- [19] Svante Janson, *Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs*, Ann. Probab. **16** (1988), no. 1, 305–312. MR-0920273

- [20] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR-1782847
- [21] Svante Janson, Krzysztof Oleszkiewicz, and Andrzej Ruciński, *Upper tails for subgraph counts in random graphs*, Israel J. Math. **142** (2004), 61–92. MR-2085711
- [22] Svante Janson and Andrzej Ruciński, *The infamous upper tail*, Random Structures Algorithms **20** (2002), no. 3, 317–342, Probabilistic methods in combinatorial optimization. MR-1900611
- [23] Svante Janson and Lutz Warnke, *The lower tail: Poisson approximation revisited*, Random Structures Algorithms **48** (2016), no. 2, 219–246. MR-3449596
- [24] Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtěch Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), no. 2, 133–163. MR-1379396
- [25] Hiroshi Maehara, *On random simplices in product distributions*, J. Appl. Probab. **17** (1980), 553–558. MR-0568966
- [26] Brendan D. McKay, *Asymptotics for symmetric 0-1 matrices with prescribed row sums*, Ars Combin. **19** (1985), no. A, 15–25. MR-0790916
- [27] Vladimir G. Mikhaïlov, *On a theorem of Janson*, Teor. Veroyatnost. i Primenen. **36** (1991), no. 1, 168–170. MR-1109028
- [28] Frank Mousset, Andreas Noever, Konstantinos Panagiotou, and Wojciech Samotij, *On the probability of nonexistence in binomial subsets*, arXiv:1711.06216 (2017).
- [29] Klaus F. Roth, *Sur quelques ensembles d’entiers*, C.R. Acad. Sci Paris **234** (1952), 388–390. MR-0046374
- [30] Andrzej Ruciński, *When are small subgraphs of a random graph normally distributed?*, Prob. Th. Rel. Fields **78** (1988), no. 1, 1–10. MR-0940863
- [31] Juanjo Rué, Christoph Spiegel, Ana Zumalacárregui, *Threshold functions and Poisson convergence for systems of equations in random sets*, A. Math. Z. **288** (2018), no. 1-2, 333–360. MR-3774416
- [32] SageMath, *The sage mathematics software system (version 8.3) [open source software package]*, www.sagemath.org (2018).
- [33] David Saxton, Andrew Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), 925–992. MR-3385638
- [34] Mathias Schacht, *Extremal results for discrete random structures*, Ann. of Math. **184** (2016), 333–365. MR-3548528
- [35] Endre Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriï Vladimirovič Linnik. MR-0369312
- [36] Bartel L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw. Arch. Wisk. **15** (1927), 212–216.
- [37] Lutz Warnke, *Upper tails for arithmetic progressions in random subsets*, Israel J. Math. **221** (2017), no. 1, 317–365. MR-3705856

Acknowledgments. The authors are grateful to the anonymous referee for constructive and helpful remarks that have improved the presentation of the paper.