Small Cliques in Random Graphs

A. D. Barbour  
Institut für Angewandte Mathematik, Universität Zürich, Rämistrasse 74,  
CH-8001 Zürich, Switzerland

Svante Janson  
Department of Mathematics, Uppsala University, Thunbergsvägen 3,  
S-752 38 Uppsala, Sweden

Michał Karoński and Andrzej Ruciński  
Department of Discrete Mathematics, Adam Mickiewicz University,  
ul. Matejki 48/49, Poznań, Poland

ABSTRACT

For $d \geq 1$, a $d$-clique in a graph $G$ is a complete $d$-vertex subgraph not contained in any larger complete subgraph of $G$. We investigate the limit distribution of the number of $d$-cliques in the binomial random graph $G(n, p)$, $p = p(n)$, $n \to \infty$.

1. INTRODUCTION AND RESULTS

Let $G(n, p)$ denote the Bernoulli graph on $n$ vertices: that is, $G(n, p) = \{J_{ij}; 1 \leq i < j \leq n\}$, where the $J_{ij}$ are independent $\text{Be}(p)$ random variables. We consider the distribution of the random variable $X_n(d)$ counting the number of $d$-cliques in a realization of $G(n, p)$. A $d$-clique in a graph $G$ is a $d$-element subset $\alpha$ of vertices which induces a complete subgraph of $G$, but such that no other vertex of $G$ is joined by edges to all vertices of $\alpha$. Thus $X_n(d)$ can be expressed as a sum of indicator random variables $I_\alpha$, where

\[
I_\alpha = \begin{cases} 
1 & \text{if } \alpha \text{ is a } d\text{-clique in } G(n, p) \\
0 & \text{otherwise} 
\end{cases}
\]  

(1.1)
and $\alpha$ runs over all elements of the set $\binom{[n]}{d} = \{\alpha \subseteq [n] : |\alpha| = d\}$, where $[n]$ denotes the set $\{1, 2, \ldots, n\}$.

Clearly, there is no hope of finding the exact distribution of $X_n(d)$, so we look for approximations when $n$ is large. As a first step, note that the mean of $X_n(d)$ is given by

$$EX_n(d) = \binom{n}{d} \pi, \quad \text{where} \quad \pi = EI_n = p^{(d)}(1 - p^d)^{n-d}.$$  \hfill (1.2)

However, the $I_n$ are mutually dependent, and even the second moment is very awkward to express. Thus direct methods, such as the method of moments, commonly used in the study of random graphs, prove to be of little use here. The arguments we use therefore typically involve two steps: first finding a random variable $X^*_n$ of simpler structure which is asymptotically close enough to $X_n(d)$ and then using general theorems, developed for proving limit theorems for sums of dependent variables, to establish the behavior of $X^*_n$.

The notion of a clique is closely related to that of a complete subgraph, since a clique is a complete subgraph which is contained in no other complete subgraph, and it is natural to begin by considering the distribution of $Y_n(d)$, the number of complete subgraphs $K_d$ on $d$ vertices contained in $G(n, p)$. This problem has been well investigated, and a good description is known. Thinking of $G(n, p)$ as $p$ increases from $0$, $K_d$'s only begin to appear when $n^2 p^{d-1} \approx 1$, at which stage $Y_n(d)$ is approximately Poisson distributed. If $n^2 p^{d-1} \gg 1$, the distribution of $Y_n(d)$ is asymptotically normal, provided that $p$ is not so large that $n^2 (1-p) = O(1)$. At the upper extreme, when $n^2 (1-p) \approx 1$, only a $\text{Bi}\left(\binom{n}{d}, (1-p)\right)$-distributed number $Z_n$ of edges is absent from $G(n, p)$, and $Y_n(d) \approx \binom{n}{d} - Z_n\left(\binom{n}{d} - \frac{1}{2}\right)$. If $n^2 (1-p) \ll 1$, $Z_n$ is zero with probability essentially 1. A more precise statement of these results is as follows. Here and subsequently, we use the notation $\tilde{Z}$ to denote the standardized version $\left(\text{Var} Z\right)^{-1/2}(Z - EZ)$ of a random variable $Z$, and $d_{TV}(\mathcal{D}_1, \mathcal{D}_2)$ for the total variation distance between the distributions $\mathcal{D}_1$ and $\mathcal{D}_2$.

**Theorem A** (Erdős and Rényi [4], Schüerger [14], Ruciński [13]). Fix $d \geq 2$, let $Y_n = Y_n(d)$ and $\mu_n = EY_n$, and write $p$ for $p(n)$. Then, as $n \to \infty$,

(a) Threshold: if $n^2 p^{d-1} \to \{0\}$, then $P[Y_n = 0] \to \{1\}$

(b) Poisson approximation:

(i) if $n^2 p^{d+1} \to 0$, then $d_{TV}(\mathcal{L}(Y_n), \text{Po}(\mu_n)) \to 0$;

(ii) if $n^2 (1-p)^3 \to 0$, then

$$d_{TV}\left(\text{Bi}\left(\left\lfloor \frac{n}{d}\right\rfloor \frac{1}{4} n^2 (1-p)^2 + 1\right), \text{Bi}\left(\left\lfloor \frac{n}{d}\right\rfloor, (1-p)\right)\right) \to 0,$$

where $\lfloor x \rfloor$ denotes the integer part of $x$;

(c) Normal approximation: if $n^2 p^{d-1} \to \infty$ and $n^2 (1-p) \to \infty$, then $\tilde{Y}_n \to N(0, 1)$. 


Remark 1.1. Part (b)(ii) is not to be found in earlier work: a proof is given in Section 2.

The behavior of $X_n(d)$, the number of $d$-cliques, is naturally very similar to that of $Y_n(d)$ when $p$ is small, since then $K_d$'s are rather rare and $K_{d+1}$'s still rarer.

From the expression

$$EY_n(d) = \binom{n}{d} p^{\frac{d}{2}}$$  \hspace{1cm} (1.3)

for the expectation of $Y_n(d)$, it can be seen that $EY_n(d+1) \ll EY_n(d)$ provided that $np_d \ll 1$, and hence it is to be expected that $X_n(d) = Y_n(d)$ in this range with a threshold for existence as in Theorem A(a), Poisson approximation as in Theorem A(b)(i), and normal approximation if $n^2 p^{d-1} \gg 1$ and $np_d \ll 1$. However, for larger values of $p$, $1 = O(np^n)$, the behavior of $X_n(d)$ differs dramatically from that of $Y_n(d)$.

The expectation (1.3) of $Y_n(d)$ is an increasing function of $p$, and the second restriction $n^2 (1-p) \to \infty$ on the range of normal approximation in Theorem A(c) arises solely because, for $p$ extremely close to 1, there is little variability left in the distribution of $Y_n(d)$, since almost all edges, and hence almost all possible copies of $K_d$, are present. For $X_n(d)$, on the other hand, the expectation (1.2) reaches a maximum in the range $np_d \approx 1$, and then declines fast because of the factor $(1-p)^{n-d}$; indeed, if $p$ is such that

$$np_d = \frac{1}{2} ((d+1) \log n + (d-1) \log \log n) + \lambda_n$$  \hspace{1cm} (1.4)

where $\lambda_n \gg 1$, then $EX_n(d) \approx 1$, implying that $P[X_n(d) = 0] \sim 1$ once again. This, by analogy with random variables such as the number of isolated trees, suggests that there is a second threshold, with an associated range of $p$ for which Poisson approximation is useful, and that between the thresholds, where $EX_n(d) \gg 1$, the distribution of $X_n(d)$ should be asymptotically normal. These results can be precisely stated as follows.

Theorem B. Fix $d \geq 2$, let $X_n$ denote $X_n(d)$, $\mu_n = EX_n$, and write $p$ for $p(n)$. Then, as $n \to \infty$, we have

(a) Thresholds:

(i) if $n^2 p^{d-1} \to 0$, then $P[X_n = 0] \to 1$

(ii) if $n^2 p^{d-1} \to \infty$ and $\lambda_n \to -\infty$, then $P[X_n = 0] \to 0$

(iii) if $\lambda_n \to \infty$, then $P[X_n = 0] \to 1$,

where $\lambda_n$ is defined via (1.4).

(b) Poisson approximation:

(i) if $n^2 p^{d+1} \to 0$, then $d_{TV}(\mathcal{L}(X_n), \text{Po}(\mu_n)) \to 0$;

(ii) if $np^d = \frac{1}{2} ((d-1) \log n + \log \log n) + \theta_n$, where $\theta_n \to \infty$,

then $d_{TV}(\mathcal{L}(X_n), \text{Po}(\mu_n)) \to 0$. 

(c) Normal approximation:

\[ n^2 p^{d-1} \to \infty \text{ and } \lambda_n \to -\infty, \text{ where } \lambda_n \text{ is defined by (1.4), then } \tilde{X}_n \to N(0, 1). \]

Note that the condition in Theorem B(c) is equivalent to requiring that \( EX_n(d) \to \infty \), or that \( \text{Var } X_n(d) \to \infty \), which are clearly necessary conditions for normal approximation.

**Remark 1.2.** For \( d = 1 \), \( X_n \) is the number of isolated vertices, a quantity whose distribution has been thoroughly investigated in Barbour, Karioński, and Ruciński [2] and in Kordecki [11]; the results corresponding to Theorem B are these:

- if \( \lambda_n = np - \log n \to \{-\infty\}, P[X_n(1) = 0] \to \{0\}; \)
- if \( np \to \infty \), \( d_{TV}(\mathcal{L}(X_n(1)), \text{Po}(\mu_n)) \to 0 \), where \( \mu_n = EX_n(1) = n(1-p)^{n-1}; \)
- if \( n^2 p \to \infty \) and \( np - \log n \to -\infty \), then \( \tilde{X}_n(1) \to N(0, 1). \)

The proof of Theorem B(a) is a straightforward application of the second moment method. Theorem B(b) is proved by the Stein–Chen method, but the range of Poisson approximation obtained may not be quite the best possible. The details are to be found in Section 2. In Section 3, the simplest normal convergence result, in the range of \( p \) for which \( X_n(d) \sim Y_n(d) \), is established. Further progress depends on more sophisticated arguments. Even computing the variance of \( X_n(d) \), a vital constituent of the proofs that follow, is no simple task. Its asymptotic behavior is described in some detail in Section 4, and this makes possible the appropriate estimates of the magnitudes of the terms in the orthogonal decomposition of \( X_n(d) \) presented in Section 5, and inspired by Janson and Nowicki [9]. This decomposition is used in the proofs contained in the last three sections. In Section 6 we consider the cases for which the first projection method works. This turns out to be a special case of a more general approach presented in Section 7. However, despite all these efforts, there are still some cases not covered. For these, we use a martingale approach first developed in Janson [7, 8], and based on a theorem of Jacod and Shiryaev [6].

As just observed, different methods are used in different ranges of \( p \) to prove the asymptotic normality stated in Theorem B(iii). In order to combine these and obtain the full result, we use the standard fact that if every subsequence of \( \{\tilde{X}_n\} \) has a subsubsequence which converges in distribution to \( N(0, 1) \) then the entire sequence converges. Consequently, if we, for example, can prove that \( \tilde{X}_n \) converges to \( N(0, 1) \) under each of the assumptions \( np^d \to 0, np^d \to c, 0 < c < \infty \), and \( np^d \to \infty \), then the result holds without any assumption because every sequence \( p(n) \) has a subsequence for which one of these cases holds. We will use this argument several times without explicitly mentioning it.

2. **Thresholds and Poisson Approximation**

To establish the results on thresholds and Poisson approximation, we mostly use methods based on simple second moment estimates. The variance of \( X_n(d) \) can
be expressed, using the notation (1.1), by the exact formula (Kalbfleisch [10])

\[
\text{Var } X_n(d) = \sum_{\alpha} \sum_{\beta} \text{Cov}(I_\alpha, I_\beta) \\
= \sum_{r=0}^{d} \binom{n}{d-r, d-r, r} \\
\times \{ p^{2^{(\frac{d}{2})}(\frac{d}{2})-1}(1-2p^d+p^{2d-r})^{n-2d+r}P_{d-r} - (p^{\frac{d}{2}}(1-p^d)^{n-d})^2 \} \\
= \sum_{r=0}^{d} A_r, \tag{2.1}
\]

where $P_r$ is the probability that the maximum degree in a random bipartite graph $G(t, t, p)$ is smaller than $t$, and we set $P_0 = 1$. In (2.1), the combinatorial coefficient gives the number of ways of choosing two $d$-sets of vertices which overlap in $r$ elements, the quantity $(1-2p^d+p^{2d-r})$ is the probability that a further given vertex is not joined either to all vertices in the first $d$-set or to all in the second, and the factor $P_{d-r}$ allows for the possibility that a vertex belonging to just one of the $d$-sets may be connected to all those in the other $d$-set. The $P_r$ can be easily computed by inclusion–exclusion, giving $P_1 = 1 - p$, $P_2 = 1 - 4p^2 + 4p^3 - p^4$, and

\[
P_t = 1 - 2tp^2 + t^2p^{2t-1} + 2\binom{t}{2}p^{2t} + O(p^{3t-2}), \quad t = 3, 4, \ldots . \tag{2.2}
\]

In Section 4 we investigate the asymptotic behavior of the variance in great detail. Here, all we need is to show that if $n^2p^{d-1}$ is bounded away from 0 and $np^d = O(\log n)$, then

\[
\text{Var } X = EX + O(n^{-1}(EX)^2). \tag{2.3}
\]

To do so, we note that

\[
A_0 = \binom{n}{d, d} p^{2^{(\frac{d}{2})}(1-p^d)^{2n-4d}}P_d - (1-p^d)^{2d} = (EX)^2p^{2d-1},
\]

\[
A_r = O(n^{2^{(\frac{d}{2})}(\frac{d}{2})-1}e^{-2np^d}) = O(n^{-1}(EX)^2), \quad r = 1, \ldots, d-1
\]

\[
A_d = EX(1-\pi) = EX + O(n^{-d}(EX)^2).
\]

Using (2.3), we can now prove Part (ii) of Theorem B(a) by the second moment method; parts (i) and (iii) are direct consequences of (1.2), which implies that $EX_n(d) \to 0$ in the given ranges of $p$.

**Proof of Theorem B(a)(ii).** By Chebyshev's inequality and (2.3),

\[
P[X_n(d) = 0] \leq \text{Var } X_n(d)/\{EX_n(d)\}^2 = O\left(\frac{1}{EX} + \frac{1}{n}\right),
\]

and hence $P[X = 0] \to 0$ whenever $EX \to \infty$, proving Theorem B(a)(ii). \hfill \blacksquare
We now turn to proving the Poisson approximation results of Theorem B(b), that \( \delta_n = d_{TV}(\mathcal{L}(X_n), \text{Po}(\mu_n)) \to 0 \) if either \( n^2 p^{d+1} \to 0 \) or \( \theta_n \to \infty \), where

\[
\theta_n = np^d - \frac{1}{2}(d - 1)(\log n + \log \log n).
\]

We use the following estimate of \( \delta_n \), based on the Stein–Chen method, due to Barbour, Holst, and Janson [1, Corollary II.F.1.].

**Lemma 2.1.** Let \( (J_{e})_{e \in E} \) be independent random variables and, for each \( \alpha \in \Gamma \), \( I_\alpha \) be a function of \( (J_{e})_{e \in E} \). If, for each \( \alpha \in \Gamma \), there exists \( \Gamma_\alpha \subseteq \Gamma \) with \( \alpha \in \Gamma_\alpha \) and a partition \( E = E_1^{(\alpha)} \cup E_2^{(\alpha)} \cup E_3^{(\alpha)} \) such that

(i) \( I_\alpha \) is a function of \( (J_{e} : e \in E_1^{(\alpha)} \cup E_2^{(\alpha)}) \) and decreasing as a function of \( (J_{e} : e \in E_2^{(\alpha)}) \);

(ii) for each \( \beta \notin \Gamma_\alpha \), \( I_\beta \) is a function of \( (J_{e} : e \in E_2^{(\alpha)} \cup E_3^{(\alpha)}) \) and decreasing as a function of \( (J_{e} : e \in E_2^{(\alpha)}) \);

then, for \( W = \sum_{\alpha \in \Gamma} I_\alpha \),

\[
d_{TV}(\mathcal{L}(W), \text{Po}(EW)) \leq (EW)^{-1}(\text{Var } W - EW + 2 \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha} E(I_\alpha)E(I_\beta)).
\]

**Proof of Theorem B(b).** If \( EX_n \to 0 \), it follows immediately that \( \delta_n \to 0 \). Otherwise apply Lemma 2.1 with \( \Gamma = [\frac{n}{2}] \), \( \Gamma_\alpha = \{ \beta \in \Gamma : \beta \cap \alpha \neq \emptyset \} \), \( E_1^{(\alpha)} = [\frac{\alpha}{2}] \), \( E_2^{(\alpha)} = [\frac{n - \alpha}{2}] \). Hence, using (2.3),

\[
\delta_n \leq (EX)^{-1}\{\text{Var } X - EX + 2(EX)^2|\Gamma_\alpha|/|\Gamma|\} = O(n^{-1}EX).
\]

Thus, if either \( n^2 p^d \to 0 \) or \( \theta_n \to \infty \), \( \delta_n \to 0 \).

It remains to prove Poisson approximation when \( n^2 p^{d+1} \to 0 \) but \( n^2 p^d \) is bounded away from 0. In this range, the general estimate of Lemma 2.1 is not sharp enough, and an explicit coupling is used together with the Stein–Chen method to obtain the result. The essential estimate is that

\[
d_{TV}(\mathcal{L}(X_n), \text{Po}(\mu_n)) \leq E|U_n - V_n|,
\]

(2.4)

for any pair of random variables on the same probability space such that \( \mathcal{L}(U_n) = \mathcal{L}(X_n) \) and \( \mathcal{L}(V_n) = \mathcal{L}(X_n - I_n, I_n = 1) \), for any given \( \alpha \in \Gamma \): this is a consequence of Barbour, Holst, and Janson [1, Theorem II.A] and symmetry.

To construct \( U_n \) and \( V_n \), realize \( G(n, p) \) as usual, and set \( U_n = X_n \). If \( I_\alpha = 1 \), set \( V_n = X_n - 1 \); otherwise define a new graph \( G^* = \{J^*_e : e \in [\frac{n}{2}]\} \) and set \( V_n = \#\{d - \text{cliques} \neq \emptyset \text{ in } G^*\} \), where the \( J^*_e \) are determined as follows. For \( k \neq \alpha \), let \( N_k = \{\{ik\}, i \in \alpha\} \), and write \( N_k = \#\{e \in N_k : J_e = 1\} \). Then

(i) If \( e \subset \alpha^c \), set \( J^*_e = J_e \);

(ii) If \( e \subset \alpha \), set \( J^*_e = 1 \);

(iii) If \( k \notin \alpha \), \( N_k < d \), and \( i \in \alpha \), set \( J^*_{ik} = J_{ik} \).
(iv) If \( k \not\in \alpha \) and \( N_k = d \), resample from the distribution \( \text{Bi}(d, p) \) until a value \( N_k^* < d \) is obtained, and define \( N_k^* \) elements of \( \{ J_e : e \in \mathcal{N}_k^* \} \) to be 1 and the remainder 0, uniformly at random.

Because of the independence of the edge sets indexed by \( [\alpha] \), \([\alpha^c]\) and the \( (\mathcal{N}_k^c) \), \( k \not\in \alpha \), this construction yields the required distribution for \( V_n \), since \( G^* \) itself is distributed as \( G \) conditional on \( I_\alpha = 1 \).

Clearly, if \( I_\alpha = 1 \),
\[
|U_n - V_n| = 1.
\] (2.5)

If not, the value of \( V_n \) can differ from that of \( X_n \) through (ii) and (iv); from (ii), because of those indices \( e \in [\alpha] \) where \( J_e = 0 \), and hence \( J_e < J^*_e = 1 \), and from (iv) because of those indices \( k \not\in \alpha \) where \( N_k = d \), and hence, for some or all \( i \in \alpha, 0 = J^*_i < J_i = 1 \).

From (ii), we define the random variables
\[
\Delta_1 = \# \{ d - \text{cliques} \neq \alpha \text{ in } G^* \text{ containing edges } e \text{ where } J^*_e > J_e \},
\]
\[
\Delta_2 = \# \{ d - \text{cliques in } G \text{ which are contained in a } K_{d+1} \text{ in } G^* \},
\]

and from (iv) the random variables
\[
\Delta_3 = \# \{ d - \text{cliques} \neq \alpha \text{ in } G^* \text{ which are contained in a } K_{d+1} \text{ in } G \},
\]
\[
\Delta_4 = \# \{ d - \text{cliques in } G \text{ containing edges } e \text{ where } J_e > J^*_e \}.
\]

Note that, if \( I_\alpha = 0 \), then
\[
|U_n - V_n| = |\Delta_2 + \Delta_4 - \Delta_1 - \Delta_3| \leq \Delta_1 + \Delta_2 + \Delta_3 + \Delta_4.
\]

Together with (2.5), this yields
\[
E|U_n - V_n| \leq E(\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4) + \pi.
\] (2.6)

In order to estimate \( E\Delta_1 \), we estimate the contributions of those \( K_r \)'s in \( G^* \) which have \( r \) vertices in \( \alpha, 2 \leq r \leq d - 1 \), and obtain
\[
E\Delta_1 = O\left( \sum_{r=2}^{d-1} n^{d-r} p^{(\frac{d}{2})-r} \right) = O(n^{d-2} p^{(\frac{d}{2})-1}) = O((n^2 p^{d+1})^{\frac{d-2}{2}}) = o(1),
\]
when \( d > 2 \), and \( \Delta_1 = 0 \) when \( d = 2 \). For \( E\Delta_2 \), a similar decomposition gives
\[
E\Delta_2 = O\left( \sum_{r=1}^{d-1} n^{d-r} p^{(\frac{d}{2})+d-r} \right) = O(n^{d-1} p^{(\frac{d}{2})+d-1}) = O((n^2 p^{d+2})^{\frac{d-1}{2}}) = o(1).
\]

Turning to the effects of (iv), suppose \( N_k = d \), then, considering \( K_{d+1} \)'s in \( G \) with \( 1 \leq r \leq d \) edges in common with \( \mathcal{N}_k^* \), we have a contribution to \( E\Delta_3 \) of \( O(\sum_{r=1}^{d} n^{d-r} p^{(\frac{d}{2})-r}) = O(n^{d-1} p^{(\frac{d}{2})-1}) \). Hence, since the expected number of
\( k \in \alpha^e \) with \( N_k = d \) is of order \( np^d \), we find that
\[
E \Delta_3 = O(n^d p^{(\frac{d}{2})+d-1}) = O((n^2 p^{d+1})^{d/2}p^{d-1}) = o(1) .
\]
Finally, considering \( K_d \)'s in \( G \) with \( 1 \leq r \leq d-1 \) edges in common with \( N_k \), we obtain
\[
E \Delta_4 = O \left( np^d \sum_{r=1}^{d-1} n^{d-r-1} p^{(\frac{d}{2})-r} \right) = O((n^2 p^{d+1})^{d/2}p^{d-1}) = o(1) .
\]
The remaining part of Theorem B(b) now follows immediately from (2.4) and (2.6).

In the case when \( d = 2 \) and \( \lambda_n \), defined by (1.4), tends to a finite limit, Poisson convergence was proved by Erdős, Palmer, and Robinson [5]. Note that the range where Poisson approximation is proved here for the larger values of \( p \), that is, when \( \theta_n \to \infty \), is smaller than the range in which \( \text{Var} X_n \sim EX_n \), specified by \( \nu_n \to \infty \), where \( \nu_n \) is defined in Lemma 4.1 below.

We complete this section by proving the outstanding Poisson approximation result of Theorem A.

**Proof of Theorem A(b)(ii).** Write \( Y_n = \binom{n}{d} - Z_n \binom{n-2}{d-2} + R_n \binom{n-2}{d-2} \). While \( R_n \) can in principle be explicitly determined by inclusion–exclusion, we instead observe that, writing \( q = 1 - p \),
\[
E R_n = EZ_n + \left( EY_n - \binom{n}{d} \right)/\binom{n-2}{d-2} = \binom{n}{2} q + \binom{n}{d} \binom{n-2}{d-2}^{-1} (1-q)^{\binom{d}{2}}-1
\]
\[
= \frac{n^2}{2} \frac{(\frac{d}{2}-1)}{2} q^2 + O(nq^2, n^2q^3) ,
\]
and
\[
\text{Var} R_n = \binom{n-2}{d-2}^{-2} \text{Var} Y_n + 2(\binom{n-2}{d-2})^{-1} \text{Cov}(Y_n, Z_n) + \text{Var} Z_n .
\]
Now the constituents of this latter expression can be evaluated as
\[
\text{Var} Y_n = \sum_{r=2}^{d} \frac{(n)^{2d-r}}{r!(d-r)!^2} (1-q)^{2\binom{r}{2}-\binom{d}{2}} \{1 - (1-q)^{\binom{d}{2}}\}
\]
\[
= \frac{(n)^{2d-2}}{2(d-2)!^2} (1-q)^{2\binom{d}{2}-1} q + \frac{(n)^{2d-4}}{3!(d-3)!^2} (1-q)^{2\binom{d}{2}-3} \{1 - (1-q)^3\}
\]
\[
+ O(n^{2d-4}q) ,
\]
\[
\text{Var} Z_n = \binom{n}{d} q(1-q) ,
\]
and
\[
\text{Cov}(Y_n, Z_n) = E(Y_n Z_n) - EY_n EZ_n
\]
\[
= \binom{n}{2} q \binom{n}{d} - \binom{n-2}{d-2} (1-q)^{\binom{d}{2}} - \binom{n}{2} q \binom{n}{d} (1-q)^{\binom{d}{2}}
\]
\[
= - \binom{n}{2} \frac{n-2}{d-2} q(1-q)^{\binom{d}{2}} .
\]
Hence

\[
\text{Var } R_n = \frac{n!(n-d)!^2q}{2(n-2d+2)!(n-2)!^2} (1-q)^{2(\frac{d}{2})-1} + \frac{n^{2d-3}q^2}{2(d-3)!^2} \left( \frac{d-2}{n^{2d-4}} - 2 \binom{n}{2} q(1-q)^{\frac{d}{2}} + \binom{n}{2} q(1-q)^{\frac{d}{2}} + O(q, nq^2) \right) \\
= \left( \binom{n}{2} q \right) \left( \frac{n-d}{n-2} \right)^2 \left( 1 - (2\binom{d}{2} - 1)q \right) + \frac{(d-2)^2}{n} - 2(1-\binom{d}{2}q) \\
+ (1-q) \right) + O(q, n^2q^3) \\
= \left( \binom{n}{2} q \right) \left( (1 - (2\binom{d}{2} - 1)q) \left( \prod_{j=2}^{d-1} \left( 1 - \frac{d-2}{n-j} \right) - 1 \right) + \frac{(d-2)^2}{n} \right) + O(q, n^2q^3) \\
= O(q, n^2q^3),
\]

and so, from Chebyshev's inequality,

\[
P \left[ \left( \frac{n-2}{d-2} \right)^{-1} \left( \binom{n}{2} - Y_n \right) + \frac{n^2q^2}{4} \left( \binom{d}{2} - 1 \right) + \frac{1}{2} \neq Z_n \right] = O(q, n^2q^3)
\]
as required.

3. NORMAL APPROXIMATION FOR SMALL $p$

As observed in the introduction, $EY_n(d+1) = o(EY_n(d))$ when $np^d = o(1)$, and it is thus plausible that, for the purposes of normal approximation, $X_n(d)$ and $Y_n(d)$ are equivalent in this range of $p$. In the following proposition, we show how to make the idea precise.

**Proposition 3.1.** If $np^d \to 0$, $\tilde{X}_n(d) \to N(0, 1)$.

**Proof.** Observe that $X = X_n(d)$ may be expressed by inclusion–exclusion as

\[
X = Z_0 - (d+1)Z_1 + \sum_{r=2}^{n-d} (-1)^r Z_r = Z_0 + R,
\]

where $Z_r$ denotes the number of copies of $M_{d,r}$ in $G(n, p)$, and the multistar $M_{d,r}$ is the graph obtained from $K_d$ by adding $r$ vertices and joining each of them to all vertices of $K_d$. Now $\tilde{Z}_0 = \tilde{Y}_n(d) \to N(0, 1)$ by Theorem A(c). Thus it is enough to prove that $\text{Var } R = o(\text{Var } \tilde{Z}_0)$.

By Minkowski's inequality, writing $\sigma_r^2 = \text{Var } Z_r$,

\[
\text{Var } R \leq \left\{ (d+1)\sigma_1 + \sum_{r=2}^{n-d} \sigma_r \right\}^2.
\]

Note that the graph $M_{d,r}$ has $v(r) = d + r$ vertices and $f(r) = (\frac{d}{2}) + dr$ edges, and
that for each subgraph $H$ of $M_{d,r}$ with at least 3 vertices

$$
(e(H) - 1)/(v(H) - 2) \leq (f(r) - 1)/(v(r) - 2),
$$

(3.3)

where $e(H)$ and $v(H)$ are the numbers of edges and vertices of $H$. Thus, from Ruciński [13], Section 2, the minimum $\psi = \min n^{v(H)} p^{e(H)}$ is achieved by $H = M_{d,r}$ when $n^{v(r) - 2} p^{f(r) - 1} \leq 1$ and by $H = K_2$ otherwise.

By a general formula (cf. Ruciński [13])

$$
\sigma_r^2 \leq \sum_{\alpha, \beta} p^{2f(\alpha) - e(\alpha)} \leq \sum_{\alpha, \beta} n^{v(\alpha)} p^{2f(\alpha) / \psi},
$$

(3.4)

where the sums run over all pairs $\alpha, \beta$ of copies of $M_{d,r}$ in $K_n$ such that their intersection $H$ has at least one edge. Let $N_j$ be the number of pairs $\alpha, \beta$ which intersect in $j$ vertices. Then, if $r \geq 2$,

$$
N_j = \binom{n}{d + r - j} \binom{d + r}{d}^2 = (n)_{2d + 2r - j} (d + r) (d + r) \binom{1}{d! r!}^2
\leq n^{2d + 2r - j} (d + r)^2 (r!)^{-2},
$$

since there are $\binom{d + r}{d}$ copies of $M_{d,r}$ on $d + r$ given vertices; if $r = 1$ we obtain $N_j = (d + r - j, d + r - j, j) \leq n^{2d + 2r - j}$. Hence,

$$
\sigma_r^2 \leq \sum_{j=0}^{d+r} N_j n^j p^{2f(r) / \psi} \leq \sum_{j=0}^{d+r} \binom{d + r}{j} (d + r)^2 (r!)^{-2} n^{2d + 2r - j} p^{2f(r) / \psi}
= (r + d + 1)^{-r+d} (r!)^{-2} n^{2v(r)} p^{2f(r) / \psi}.
$$

(3.5)

Since $(r + d + 1)^{-r+d} = o((r!)^2)$ as $r \to \infty$, and $np^d \to 0$, we obtain

$$
\sigma_r^2 \leq C_1 n^{2v(r)} p^{2f(r) / \psi} = C_1 \max(n^{v(r)} p^{f(r)}, n^{2v(r) - 2} p^{2f(r) - 1})
= C_1 \max((np^d)^{v(0)} p^{f(0)}, (np^d)^{2r} n^{2v(0) - 2} p^{2f(0) - 1})
\leq C_2 (np^d)^2 \sigma_0^2,
$$

(3.6)

where $C_1$ and $C_2$ are independent of both $n$ and $r$ (see again Ruciński [13] for the asymptotic behaviour of $\sigma_0^2$).

Hence

$$
\text{Var} \frac{R}{\sigma_0^2} = O\left(\sum_{r=1}^{\infty} (np^d)^{-r/2}\right)^2 = O(np^d) = o(1).
$$

4. THE ASYMPTOTICS OF VARIANCE

A precise analysis of the asymptotic behavior of the variance of $X = X_n(d)$ will be crucial for our further arguments. Recall from Section 2 that
\[ \text{Var } X = \sum_{r=0}^{d} A_r^{(d)} , \]

where \( A_r = A_r^{(d)} \) is the sum of covariances of those pairs \((I_x, I_{-x})\) whose indices intersect on an \( r \)-element set. Let us observe that if \( np^d = O(\log n) \) then, for \( d \geq 2 \),

\[ A_r^{(d)} \sim \frac{n^{2d-r} \pi^2}{(d-r)!^2 r!} p^{-(r)} , \quad r = 2, \ldots, d-1 \quad (4.1) \]

and

\[ A_d^{(d)} \sim \frac{n^d}{d!} \pi \sim EX , \quad (4.2) \]

where \( \pi = p^{(2)}(1-p^d)^{n-d} \).

With some more effort one can show that

\[ A_0^{(d)} \sim \frac{n^{2d} \pi^2}{(d-1)!^2} p^{2d-1} \quad (4.3) \]

and

\[ A_1^{(d)} = \begin{cases} 
\frac{n^{2d-1} \pi^2}{(d-1)!^2} ((x-2(d-1)+\alpha(1))p^{d-1} & \text{when } d \geq 3 , \\
n^3 \pi^2 (x-1+\alpha(1))p & \text{when } d = 2 , 
\end{cases} \quad (4.4) \]

where \( x = x_n = np^d = O(\log n) \). Hence, by simple comparison, we realize that for \( d \geq 3 \)

\[ \text{Var } X \sim \begin{cases} 
A_2 + A_d & \text{if } x \to 0 \\
A_0 + A_1 + A_d & \text{if } x \to \infty 
\end{cases} \]

and that for \( d = 2 \), \( \text{Var } X \sim A_2 \sim EX \) if either \( x \to 0 \) or \( x \to \infty \). However, to find an asymptotic formula for \( \text{Var } X \) when \( x \to c > 0 \) we need much more precision. The reason is that although \( A_0 = A_2 \geq A_r, \ r \geq 3 \), and \( A_1 = O(A_0) \), \( A_1 \) may be negative, and so the leading terms may cancel. The next lemma provides overall information about the asymptotics of \( \text{Var } X \).

**Lemma 4.1.** Let \( A_r^{(d)}, \pi \) and \( x \) be as above and assume that \( x = O(\log n) \).

(a) For \( d = 2 \),

\[ \text{Var } X \sim \begin{cases} 
A_2 \sim \frac{1}{2} n^2 p & \text{if } x \to 0 \\
A_0 + A_1 + A_2 \sim \frac{1}{2} n^2 p C & \text{if } x \to \infty 
\end{cases} \]

where \( C = e^{-2c}(e^c + 4c^2 - 2c) \) is always positive.
(b) For $d \geq 3$, if $x \to 0$ and

$$np^{(d+1)/2} \to \begin{cases} 0 & \text{then} \ Var \ X \sim \begin{cases} A_d \sim \frac{1}{d!} n^d p^{(d)} \\ A_1 \sim \frac{n^d}{2(d-1)!} p^{d(d-1)-1} \end{cases} \\ \infty \end{cases},$$

if $x \to c \neq \frac{d-1}{2}$ then

$$Var \ X \sim A_0 + A_1 + A_2 \sim \frac{n^{2d-2} \pi^2}{2(d-1)! p} (2c - (d-1))^2;$$

if $x \to \frac{d-1}{2}$ then

$$Var \ X \sim A_0 + A_1 + A_2 + A_3 \sim \begin{cases} \frac{n^{2d-2} \pi^2}{2(d-1)! p} \{2x - d + 1\}^2 + \frac{1}{6} (d-1)(d+1)^2 p \}^{d-2} & \text{for } d \geq 4 \\ \frac{n^d \pi^2}{2p} \{(x-1)^2 + (2 + \frac{5}{3})p\} & \text{for } d = 3; \end{cases}$$

if $x \to \infty$ and

$$v_n \to \begin{cases} -\infty & \text{then} \ Var \ X \sim \begin{cases} A_0 + A_1 + A_d \sim \frac{2n^{2d}\pi^2}{(d-1)! p} p^{2d-1} \\ A_d \sim \frac{1}{d!} n^d \pi \end{cases} \\ c \end{cases},$$

where $a = \frac{2}{(d-1)!} \left( \frac{(d-1)(d-2)}{2d} \right)^{d+1-1/d} e^{-2c}$, $b = \frac{1}{d!} \left( \frac{(d-1)(d-2)}{2d} \right)^{d-1/2} e^{-c}$, and

$$v_n = np^d - \frac{1}{2d} \left((d-1)(d-2) \log n + (d^2 + 3d - 2) \log \log n \right).$$

\textbf{Proof.} All cases except when $d \geq 3$ and $x \to (d-1)/2$ follow easily from (4.1)–(4.4). In this exceptional case all we need are the following more refined estimates, valid for $x = O(1)$:

$$A_0^{(3)} = \frac{1}{4} \left( n_0 \pi^2 (p^5 - p^6 + O(p^7)) \right), \quad (4.5)$$

$$A_1^{(3)} = \frac{1}{4} \left( n_5 \pi^2 ((x-4)p^2 - (x-8)p^3 + O(p^4)) \right), \quad (4.6)$$

$$A_2^{(3)} = \frac{1}{2} \left( n_4 \pi^2 p^{-1} (1 + (x-2)p + O(p^2)) \right), \quad (4.7)$$

and, for $d \geq 4$,

$$A_0^{(d)} = \frac{(n_2 d \pi^2}{(d-1)!^2} \left( p^{2d-1} - p^{2d} + O(p^{3d-2}) \right), \quad (4.8)$$

$$A_1^{(d)} = \frac{(n_2 d-1 \pi^2}{(d-1)!^2} \left( (x-2d+2)(p^{d-1} - p^d) + (d-1)^2 p^{2d-3} + O(p^{2d-2}) \right), \quad (4.9)$$
\[ A_2^{(d)} = \frac{(n)_{2d-2} \pi^2}{2(d-2)!^2 p} (1 - p + (x - 2d + 4)p^{d-2} + O(p^{d-1})) . \] 

(4.10)

Note that (4.5), (4.6) are special cases of (4.8), (4.9), but that (4.10) does not hold for \( d = 3 \). We sketch the proof of (4.5)–(4.10), leaving the details to the reader. We have

\[
A_0 = \frac{1}{d!^2} \left( n \right)_{2d} \pi^2 (1 - p^d)^{-2d} (P_d - (1 - p^d)^{2d})
\]

\[
= \frac{1}{d!^2} \left( n \right)_{2d} \pi^2 (1 + O(p^d))(d^2 p^{2d-1} - d^2 p^d + O(p^{3d-2})) ,
\]

which gives (4.5) and (4.8). To get the other estimates we write

\[ A_1 = \frac{(n)_{2d-1}}{(d-1)!^2 \pi^2} (e^{N_1} - 1) \]

and

\[ A_2 = \frac{(n)_{2d-2} \pi^2}{2(d-2)!^2 p} (e^{N_2} - p) , \]

where

\[ N_1 = (n - 2d + 1) \log(1 - 2p^d + p^{2d-1}) + \log P_{d-1} - 2(n - d) \log(1 - p^d) \]

and

\[ N_2 = (n - 2d + 2) \log(1 - 2p^d + p^{2d-2}) + \log P_{d-2} - 2(n - d) \log(1 - p^d) , \]

and make use of the expansions \( \log(1 - y) = -y - \frac{y^2}{2} + O(y^3) \) and \( e^y = 1 + y + O(y^2) \).

Having proved (4.5)–(4.10) we see that for \( d \geq 4 \) and \( x \approx 1 \),

\[ A_0 + A_1 + A_2 = \frac{n^{2d-3} \pi^2}{2(d-1)!^2 p} (1 - p) \]

\[ \times \left\{ (2x - d + 1)^2 + (d - 1)^2 (3x - 2d + 4)p^{d-2} + O(p^{d-1}) \right\} . \]

If \( |x - \frac{d-1}{2}| = O(p^{(d-2)/2}) \) then the second term becomes leading and we must take \( A_3 \) into account, since

\[ A_3 \sim \frac{n^{2d-3} \pi^2}{(d-3)!^2 6p^3} = \frac{n^{2d-3} \pi^2}{2(d - 1)!^2 p} \frac{(d - 1)^2 p^{d-2} (d - 2)^2}{3x} \]

is of the same order. The result now follows by a simple calculation. The case \( d = 3 \) is similar. \( \square \)

**Remark 4.1.** For \( d = 1 \),

\[
\text{Var } X \sim \begin{cases} 
A_0 + A_1 \sim 2nx & \text{if } x \to 0 \\
A_0 + A_1 \sim ne^{-2x}(e^x - 1) + x & \text{if } x \approx 1 \\
A_1 \sim ne^{-x} & \text{if } x \to \infty, x = o(n^{1/2}) . 
\end{cases}
\]
5. ORTHOGONAL DECOMPOSITION

Let $G$ be a graph on the vertex set $[v]$ with $e(G)$ edges, and let $[n]_v$ be the set of all $v$-element sequences of distinct elements of $[n]$. For $\alpha = (x_1, \ldots, x_v) \in [n]_v$, let $G_\alpha$ be the copy of $G$ on the vertex set $\alpha$ under the isomorphism $i \mapsto x_i$, $i = 1, \ldots, v$. Recall that $J_e$ is the indicator of the presence of the edge $e$, $e \in [\frac{n}{2}]$, and set

$$J'_e = J_e - p.$$  

We define

$$S(G) = S_\alpha(G) = \sum_{\alpha \in [n]_v} \prod_{e \in G_\alpha} J'_e.$$  

It follows that if $G$ has $k$ isolated vertices and $\tilde{G}$ is what is left after removing them from $G$ then

$$S(G) = (n - v + k)_k S(\tilde{G}).$$

If $G$ is an empty graph, i.e., $e(G) = 0$, then we set $\tilde{G} = \emptyset$ and $S(\emptyset) = 1$.

The random variables $S(G)$ have several nice properties. First of all, $ES(G) = 0$ whenever $\tilde{G} \neq \emptyset$. Moreover, if $G = \tilde{G}$ then

$$\text{Var} S(G) = \text{aut}(G)(n)_v [p(1 - p)]^{e(G)},$$

where aut$(G)$ is the number of automorphisms of $G$. But the most important is that, if $G_1$ and $G_2$ are not isomorphic and neither has isolated vertices, then

$$\text{Cov}(S(G_1), S(G_2)) = 0. \quad (5.1)$$

Our goal is to express $X$ as a linear combination of $S(G)$'s and then to find the terms with variance of leading order. We have

$$X = \sum_{\alpha \in [\frac{n}{2}]} I_\alpha,$$

and

$$I_\alpha = \prod_{e \in [\frac{n}{2}]} (J'_e + p) \prod_{r \in \alpha} \left(1 - \prod_{s \in \alpha} (J'_{rs} + p)\right).$$

The first product can be rewritten as

$$\sum_{H} p^{(\frac{n}{2}) - e(H)} \prod_{e \in H} J'_e,$$

where $\Sigma_H$ is taken over all graphs with vertex set $\alpha$.

For each $r \in [n] \setminus \alpha$, the summands obtained by cross-multiplying the product $\prod (J'_{rs} + p)$ are of the form

$$p^{d - |\beta|} \prod_{s \in \beta} J'_{rs},$$
where $\beta$ is a subset of $\alpha$. Thus a single term in the expansion of $\Pi_s(1 - \Pi_s(J_n^\prime + p))$ is determined by a partition of $[n]\setminus\alpha$ into $2^d$ subsets $V_{\beta}, \beta \subseteq \alpha$, and so

$$X = \sum_{\alpha \in [n]} \sum_{H} p^{(\frac{1}{d})-\epsilon(H)} \prod_{e \in H} J_e' \sum_{(V_{\beta})_{\beta \subseteq \alpha}} (1 - p^d)^{|V_{\beta}|}(-1)^{n-d-|V_{\beta}|} \cdot \prod_{\beta \subseteq \alpha, \beta \neq \emptyset} \prod_{r \in V_{\beta}} \prod_{s \in \beta} J_{rs}'. \tag{5.2}$$

(The variables $H, V_{\beta}$ in the above expression should also be indexed by $\alpha$, which, however, is omitted for clarity.)

Still skipping the index $\alpha$, denote by $G$ the graph with vertex set $\alpha \cup \bigcup_{\beta \neq \emptyset} V_{\beta}$ and edge set

$$E(H) \cup \{(r, s) : r \in V_{\beta}, s \in \beta, \beta \subseteq \alpha\}.$$  

(\text{The graph } G \text{ depends not only on } \alpha \text{ but also on the choice of } H \text{ and } (V_{\beta})_{\beta \subseteq \alpha}. \text{ Thus, setting } v_{\beta} = |V_{\beta}|,

$$S(G) = \sum_{\gamma \in [n]} \prod_{r \in H} s_{\gamma} = \sum_{\alpha \subseteq [n]} \sum_{(V_{\beta})_{\beta \subseteq \alpha}} \prod_{\alpha \subseteq [n]} v_{\beta}^{-1} \prod_{e \in H} J_e' \prod_{\beta \subseteq \alpha} \prod_{r \in V_{\beta}} \prod_{s \in \beta} J_{rs}'. \tag{5.3}$$

where the second sum is taken over all partitions of the set $[n]\setminus\alpha$ into $\bigcup_{\beta \subseteq \alpha} V_{\beta}$ with $|V_{\beta}| = v_{\beta}$.

Actually, $S(G)$ depends only on the isomorphism type of $H$ (in the sense described at the beginning of this section) and on the partition of the number $n-d$ into $\Sigma_{\beta \subseteq [d]} v_{\beta}$, and therefore, by comparing (5.2) and (5.3), we arrive at the following result.

**Lemma 5.1.**

$$X = \frac{1}{d!} \sum_{H} \sum_{(V_{\beta})_{\beta \subseteq [d]}} a_{G} S(G)$$

$$= \frac{1}{d!} \sum_{H} \sum_{(V_{\beta})_{\beta \subseteq [d]}} a_{G} \left(n-d - \sum_{\beta \neq \emptyset} v_{\beta} + k\right) S(G), \tag{5.4}$$

where

$$a_{G} = p^{(\frac{1}{d})-\epsilon(H)}(1-p^d)^{v_0}(-1)^{n-d-v_0}p^{v_{\beta \neq \emptyset}(-1)^{|\beta|-v_{\beta}}} / \prod_{\beta \neq \emptyset} v_{\beta}!$$

$k$ is the number of isolated vertices in $G$, the first sum is taken over all graphs on vertex set $[d]$, the second sum is taken over all partitions of $n-d$ into nonnegative integers $v_{\beta}$, $\beta \subseteq [d]$, and the graph $G$ has vertex set $[d] \cup \bigcup_{\beta \neq \emptyset} V_{\beta}^0$ and edge set $E(H) \cup \{(r, s) : r \in V_{\beta}^0, s \in \beta, \beta \subseteq [d]\}$, where $(V_{\beta})_{\beta \subseteq [d]}$ is any partition of $[n][d]$ with $|V_{\beta}^0| = v_{\beta}, \beta \subseteq [d]$. \[\blacksquare\]
Note that only vertices from \([d]\) may be isolated in \(G\) and that the term corresponding to the empty \(\tilde{G}\) (obtained from an empty \(H\) and all \(\nu_\beta = 0\), \(\beta \neq \emptyset\)) is just \(EX\).

It may happen that different choices of \(H\) and \((\nu_\beta)\) give isomorphic (in usual sense) graphs \(\tilde{G}\) and thus the same \(S(\tilde{G})\). The orthogonal expansion of \(X\) is therefore obtained by rearranging the sum in (5.4) and combining all terms with isomorphic \(\tilde{G}\).

The exact expansion (5.4) is too complicated for our purposes and we will reduce it to the simpler approximate expansions given in Lemma 5.2 below. As an intuitive motivation for the choice of terms that we use, notice that if \(a_G S(G)\) and \(a_{G_1} S(G_1)\) are two terms in (5.4) such that \(G\) and \(G_1\) have the same vertices and the same isolated vertices and \(G \subseteq G_1\), then

\[
\text{Var}(a_G S(G)) \approx p^{|E(G)| - |E(G_1)|} \text{Var}(a_{G_1} S(G_1)).
\]

Consequently, it is reasonable to leave only those terms, where \(H\) is the complete graph on a vertex set \(\delta \subseteq [d]\) plus isolated vertices on \([d]\)\(\setminus \delta\), and \(\nu_\beta = 0\) whenever \(\beta \neq \emptyset, \delta\). It can similarly be seen that some of these terms also have comparatively small variance, which leads us to consider an even narrower class of terms.

To describe this class of graphs we recall the multistar \(M_{d,k}\) defined in Section 3. Note that \(M_{d,1} = K_{d+1}\). As a further preparation for the forthcoming lemma we define

\[
a_r = \frac{(n - r)(d - r)}{r!(d - r + 1)!} \left(\frac{1}{p^d}ight)^{(1 - p^d)(n - d - \frac{r}{2} - 1)} ((d - r + 1)(1 - p^d) - r(n - d)p^d) ,
\]

(5.5)

for \(r = 2, \ldots, d\), and

\[
b_k = \frac{(-1)^k}{d!k!} (1 - p^d)^{n - d - k}, \quad k = 1, 2, \ldots, n - d.
\]

(5.6)

**Lemma 5.2.** With notation as above and under the assumption that \(np^d = O(\log n)\), we have

\[
X - EX = \sum_{r=2}^{d} a_r S(K_r) + \sum_{k=1}^{n-d} b_k S(M_{d,k}) + R_0,
\]

where all the terms on the right-hand side are mutually orthogonal, have mean zero and, moreover, \(ER_i^2 = o(\text{Var} \, X)\). This can be in many cases further simplified as follows, where \(ER_i^2 = o(\text{Var} \, X)\) for each \(i\).

(a) For \(d = 2\),

if \(np^2 \rightarrow 0\) then

\[
X - EX = \begin{cases} a_2 S(K_2) + R_1, & \text{if } n^{-1/2} \rightarrow 0, \\ \sum_{k=2}^{n-2} b_k S(M_{2,k}) + R_2, & \text{if } n^{-1/2} \rightarrow \infty. \end{cases}
\]
(b) For \( d \geq 3 \), if \( np^d \to 0 \) and
\[
\frac{np^{(d+1)/2}}{c} \to_{\infty} 0 \quad \text{then} \quad X - EX = \begin{cases} 
 a_d S(K_d) + R_3 \\
 a_2 S(K_2) + a_d S(K_d) + R_4 \\
 a_2 S(K_2) + R_5; 
\end{cases}
\]
if \( np^d \to c \neq \frac{d-1}{2} \) then
\[X - EX = a_2 S(K_2) + R_6;\]
if \( np^d \to \frac{d-1}{2} \) and \( d \geq 4 \) then
\[X - EX = a_2 S(K_2) + a_3 S(K_3) + R_7;\]
but the term \( a_2 S(K_2) \) can be dropped when \( (np^d - \frac{d-1}{2})^2 \gg p^{d-2} \);
if \( d = 3 \), \( np^3 \to 1 \) and \( (np^3 - 1)^2 \gg p \) then
\[X - EX = a_2 S(K_2) + R_8;\]
if \( np^d \to \infty \) and
\[
\nu_n \to \begin{cases} 
 -\infty \\
 c \\
 \infty 
\end{cases} \quad \text{then} \quad X - EX = \begin{cases} 
 a_2 S(K_2) + R_9 \\
 a_2 S(K_2) + \sum_{k=2}^{n-d} b_k S(M_{d,k}) + R_{10} \\
 \sum_{k=2}^{n-d} b_k S(M_{d,k}) + R_{11}, 
\end{cases}
\]
where \( \nu_n \) is defined in Lemma 4.1.

**Proof.** Since the number of terms in (5.4) grows with \( n \), and, furthermore, some of them may coincide or cancel, it would be rather complicated to prove the lemma by estimating the variances of omitted terms. Instead, after performing all possible reductions, we will show that the variances of leading terms equal approximately \( \text{Var} X \).

Consider first the terms in (5.4) with \( \tilde{G} = K_r, 2 \leq r \leq d \). There are two groups of such terms for each \( r \): we may either select \( \delta \subseteq [d] \), let \( H \) be complete on \( \delta \) (plus the remaining vertices isolated) and set \( \nu_{\beta} = 0 \) for all \( \beta \neq \emptyset \); or we may select \( \delta \subseteq [\bar{r} + 1, d] \), let \( H \) be complete on \( \delta \) and set \( \nu_{\emptyset} = 1 \) and \( \nu_{\beta} = 0, \beta \neq \emptyset, \emptyset \). Together these terms give a contribution of
\[
\frac{1}{d!} \binom{d}{r} p^{(\ell^d - (\ell^d)} (1 - p^d)^{n-d}(n-r)_{d-r} S(K_r)
\]
\[
+ \frac{1}{d!} \binom{d}{r-1} p^{(\ell^d - (\ell^d+1)} (1 - p^d)^{n-d-1}(-1)p^{d-r+1}(n-r)_{d-r+1} S(K_r) = \alpha_r S(K_r).
\]

We next consider the terms with \( G = \tilde{G} = M_{d,k}, k \geq 1 \). For each \( k \leq n - d \) there is only one such term, obtained by taking \( H = K_d, \nu_{[d]} = k \) and \( \nu_{\beta} = 0 \) for all \( \beta \neq \emptyset, [d] \). This term is just \( b_k S(M_{d,k}) \). Now it is only a matter of tedious calculations to show that for each \( i = 0, \ldots, 11 \), \( \text{Var}(X - R_i) \) has the same
asymptotics as that of $\text{Var}(X)$ given in Lemma 4.1, and thus

$$\text{Var}(R_i) = \text{Var} X - \text{Var}(X - R_i) = o(\text{Var} X).$$

For example, if $d = 3$ and $np^3 = x \to 1$, then

$$\text{Var}(a_2 S(K_2)) \sim \frac{n^2 p^4}{16} (1 - p^3)^2 (1 - 2np^3) 2n^2 p^3$$

$$= \frac{n^4 \pi^2}{2p} (x - 1)^2 + o(p^3),$$

$$\text{Var}(a_3 S(K_3)) \sim \frac{1}{36} (1 - p^3)^2 (1 - 3p^3 - 3n^3 p^3) 2n^3 p^3$$

$$\sim \frac{n^4 \pi^2}{6x} (1 - 3x + 8p^3)^2 \sim \frac{2}{3} n^4 \pi^2,$$

$$\text{Var}(b_1 S(M_{3,1})) \sim \frac{1}{36} (1 - p^3)^2 24n^4 p^6 \sim \frac{2}{3} n^4 \pi^2,$$

$$\text{Var} \left( \sum_{k=2}^{n-3} b_k S(M_{3,k}) \right) = \sum_{k=2}^{n-3} \frac{1}{6k!} (1 - p^3)^2 (n-3-k)(n-3)_k [p(1-p)]^{3+3k}$$

$$\sim \left( \frac{n}{3} \right) p^3 (1 - p^3)^2 (n-3-k) \sum_{k=2}^{n-3} \left( \frac{n-3}{k} \right) [p(1-p)]^{3k} (1 - p^3)^{-2k}$$

$$= \left( \frac{n}{3} \right) p^3 \pi^2 \left\{ \left( 1 + \frac{p(1-p)}{(1-p^3)^{3/2}} \right)^{n-3} - 1 - (n-3) \left( \frac{p(1-p)^3}{(1-p^3)^2} \right) \right\}$$

$$= \left( \frac{n}{3} \right) \frac{n}{x} \pi^2 (1 + p^3 + O(p^4))^{n-3} - 1 - x + O(p)$$

$$\sim \frac{n^4 \pi^2}{6} (e - 2).$$

Together we obtain in that case

$$\text{Var}(X - R_0) \sim \frac{n^4 \pi^2}{6p} \{3(x-1)^2 + 4p + 4p + (e-2)p\}$$

$$= \frac{n^4 \pi^2}{2p} \left\{ (x-1)^2 + \left( 2 + \frac{e}{3} \right) p \right\}$$

which is asymptotically the same as $\text{Var} X$; see Lemma 4.1.

The verification of other cases is similar and therefore omitted.  

\begin{remark} \text{Remark 5.1.} For $d = 1$, Lemma 5.1 gives the exact decomposition $X - EX = \sum_{k=1}^{n-1} b_k S(M_{1,k})$, where $b_k$ is as in (5.6); if $np \to 0$, this can be simplified to the approximate decomposition $X - EX = b_1 S(M_{1,1}) + R = b_1 S(K_2) + R$, where $ER^2 = o(\text{Var} X)$. \end{remark}
6. THE FIRST PROJECTION

A common method for proving normality of graph statistics is the method of the first projection, which works as follows. Define, \( X^*_e = E(X - EX|J_e) \), for each \( e \in [\frac{n}{2}] \), and set

\[
X^* = \sum_{e \in [\frac{n}{2}]} X^*_e.
\]

Since \( X^*_e \) is a function of \( J_e \) which, in turn, assumes only the values of 0 and 1, and since furthermore \( EX^*_e = 0 \), \( X^*_e = c(J_e - p) \) for some constant \( c \) which may depend on \( n \) and \( p \) but not, by symmetry, on \( e \). Thus, with \( N = \sum J_e \), \( X^* = c(N - EN) \) and provided \( c \neq 0 \),

\[
\tilde{X}^* = \text{sign}(c) \tilde{N}.
\]

The variable \( N \) has a binomial distribution \( \text{Bi}(\binom{n}{2}, p) \), and thus \( \tilde{N} \to N(0, 1) \) provided that \( np(1 - p) \to \infty \). Consequently, \( \tilde{X}^* \to N(0, 1) \), and if

\[
\text{Var}(X - X^*) = o(\text{Var} X), \quad (6.1)
\]

then also

\[
\tilde{X} \to N(0, 1).
\]

Moreover, it is easily seen that \( \text{Cov}(X^*, X - X^*) = 0 \), and therefore condition (6.1) is equivalent to

\[
\text{Var}(X^*) \sim \text{Var}(X). \quad (6.2)
\]

It is not difficult to compute \( X^* \) directly from the definition and then check, using Lemma 4.1, for what range of \( p \) (6.2) holds. We prefer, however, to emphasize the usefulness of the orthogonal decomposition from Section 5 by observing that

\[
X^* = a_2 S(K_2). \quad (6.3)
\]

This is because, for \( m \geq 1 \) different elements \( e_1, \ldots, e_m \) of \( [\frac{n}{2}] \), \( E(J_{e_1} \ldots J_{e_m}|J_e) = 0 \) unless \( m = 1 \) and \( e_1 = e \). Therefore

\[
E(S(G)|J_e) = \begin{cases} 0 & \text{if } e(G) \geq 2 \\ 2J_e' & \text{if } G = K_2 \end{cases}
\]

and by Lemma 5.1 and (5.5), \( E(X - EX|J_e) = 2a_2 J_e' \), giving (6.3). Hence we see that the first projection yields the asymptotic normality of \( X \) in all cases when \( \text{Var}(X) \sim \text{Var}(a_2 S(K_2)) \), that is, by Lemma 5.2, in the following cases:

(a) \( d = 2 \) and \( x \to 0 \),
(b) \( d \geq 3 \) and
(i) $x \to 0$ and $np^{d+1} \to \infty$,
(ii) $x \to c \neq \frac{d-1}{2}$,
(iii) $x \to \frac{d-1}{2}$ and $(x - \frac{d-1}{2})^2 \gg p^{d-2}$,
(iv) $x \to \infty$ and $\nu_n \to -\infty$,

where $x = np^d$ and $\nu_n$ is as defined in Lemma 4.1.

7. FINITE DECOMPOSITIONS AND TRUNCATION

We saw in Section 6 that the method of first projection yields asymptotic normality of $X$ whenever $\text{Var}(X)$ is dominated by $\text{Var}(a_2 S(K_r))$. Using the following result from Janson [8], one can extend this to cases when $\text{Var}(X)$ is dominated by another term or a finite number of terms. (Here "finite" means not depending on $n$.)

**Lemma 7.1.** (Janson [8]). Let $\mathcal{H}$ be a finite family of nonisomorphic connected graphs and let $p \to 0$ in such a way that

$$np^{m(H)} \to \infty$$

(7.1)

for all $H \in \mathcal{H}$, where $m(H) = \max \{ s(F) : F \subseteq H \}$. Then the normalized variables $S(H), H \in \mathcal{H}$, converge jointly to independent standard normal variables. Hence, if $X^* = \sum_{H \in \mathcal{H}} c_n(H) S(H)$, for some constants $c_n(H)$, then $X^* \to N(0, 1)$. ■

In many cases Lemma 5.2 expresses $X - EX$ as such a finite sum plus a small remainder term. The graphs used in these decompositions are $K_r, r = 2, \ldots, d$ and $M_{d,k}, k = 1, 2, \ldots$. They are all connected and it is easily seen, by an argument similar to that used in the proof of Proposition 3.1, that

$$m(K_r) = \frac{r-1}{2} \quad \text{and} \quad m(M_{d,k}) < d,$$

and hence that (7.1) holds for all of them.

This procedure thus proves Theorem B(iii) in the cases when $d \geq 2$ and $np^d \to 0$ (the case already treated in Section 3) or $d \geq 3$ and $np^d \to c$ except when $d = 3, c = 1$ and $(np^d - 1)^2 = O(p)$.

In this exceptional case as well as when $d = 2$ and $np^2 \to c$ one can employ a truncation. In doing so we rely on Theorem 4.2 from Billingsley [3]. Let us set

$$X^*_m = \sum_{r=2}^d a_r S(K_r) + \sum_{k=1}^m b_k S(M_{d,k}),$$

$\sigma^2 = \text{Var}(X)$ and $R_m = X - EX - X^*_m$. According to that theorem it is enough to show that for $m = 1, 2, \ldots$,

$$\frac{1}{\sigma} X^*_m \to N(0, c_m) \quad \text{as} \quad n \to \infty,$$
where \( c_m \to 1 \) as \( m \to \infty \), and

\[
\lim_{m \to \infty} \lim_{n \to \infty} \frac{\text{Var}(R_m)}{\sigma^2} = 0.
\]

We know already that \( \tilde{X}_m \to N(0, 1) \) and thus all we need is a fair estimate of \( \text{Var}(R_m) \).

We have

\[
\text{Var}(R_m) = \left[ p(1 - p) \right]^{(\frac{1}{d})} \frac{(n)_d}{d!} \sum_{k=m+1}^{n-d} \binom{n-d}{k} \left[ p(1 - p) \right]^{kd} \left[ (1 - p^d)^2 \right]^{n-d-k}
\approx \frac{1}{d!} n^d p^{(\frac{1}{d})} \left[ ((p(1 - p))^d + (1 - p^d)^2)^{n-d} - \sum_{k=0}^{m} \frac{x^k}{k!} e^{-2x} \right]
\approx \frac{1}{d!} n^d p^{(\frac{1}{d})} e^{-2x} \left( e^x - \sum_{k=0}^{m} \frac{x^k}{k!} \right).
\]

Hence for \( d = 2 \) and \( x = np^2 \to c \), \( \sigma^{-2} \text{Var}(R_m) \to r_m(c) / (e^c + 4c^2 - 2c) \), whereas for \( d = 3 \), \( c = 1 \), and \( (x - 1)^2 = O(p) \), \( \sigma^{-2} \text{Var}(R_m) \to r_m(1) / [3(x - 1)^2/p + 6 + e] \), where \( r_m(c) = e^c - \sum_{k=0}^{m} c^k/k! \to 0 \) as \( m \to \infty \). In either case \( \lim_{n \to \infty} \sigma^{-2} \text{Var}(R_m) \to 0 \) as \( m \to 0 \) and so

\[
\lim_{m \to \infty} c_m = \lim_{m \to \infty} \left( \lim_{n \to \infty} \left( 1 - \frac{\text{Var}(R_m)}{\sigma^2} \right) \right) = 1,
\]

completing the proof in the above two cases.

8. A MARTINGALE METHOD

In Sections 3–7 we have proved Theorem B(iii) in all cases when \( np^d = O(1) \) and for \( d \geq 3 \) even when \( np^d \to \infty \) but sufficiently slowly, i.e., when at the same time \( \nu_n \to -\infty \). In the remaining cases, we cannot use the results of Janson [8] directly, but we can adapt the methods used there, which are based on a martingale convergence theorem by Jacod and Shiryaev [6].

We first introduce a time variable, which enables us to regard various statistics as stochastic processes. This is done as follows.

Let \( \{T_e\}, e \in \binom{\mathcal{E}}{2} \), be i.i.d. random variables with a uniform distribution on \([0, 1]\), and let \( J_e(t) = I(T_e \leq t) \). We regard these variables as the edge indicators of a random graph at time \( t \), and obtain thus a random graph that grows with time. At time \( t \), we have a random graph \( G_{n,t} \), \( 0 \leq t \leq 1 \). We also define

\[
J_e^*(t) = J_e(t) - t, \quad 0 \leq t \leq 1, \quad (8.1)
\]

\[
\dot{J}_e(t) = (1 - t)^{-1} J_e^*(t) = \frac{J_e(t) - t}{1 - t}, \quad 0 \leq t < 1. \quad (8.2)
\]

It is then easy to see that each \( \dot{J}_e(t) \), \( 0 \leq t < 1 \), is a martingale on \([0, 1]\) with respect to the \( \sigma \)-fields \( \mathcal{F}_t \), generated by \( \{J_e(s) : s \leq t\} \), since \( \dot{J}_e(t) \) is a Markov process and \( E(\dot{J}_e(t) | \mathcal{F}_s) = \dot{J}_e(s) \).
Now define
\begin{equation}
\hat{X}_n(t) = \sum_{a \subseteq \{2\}} \prod_{e \in a} \hat{J}_e(t) \prod_{i \not\in a} \left(1 - \prod_{j \in a} \hat{J}_j(t)\right), \tag{8.3}
\end{equation}
\begin{equation}
\hat{Z}_n(t) = \sum_{e \in \{2\}} \hat{J}_e(t). \tag{8.4}
\end{equation}

Note that, taking \( t = p \),
\begin{equation}
\hat{Z}_n(p) = \frac{1}{2} (1 - p)^{-1} S_n(K_2). \tag{8.5}
\end{equation}

Since \( \hat{X}_n(t) \) can be expanded into a sum of terms which are products \( \prod \hat{J}_e(t) \) of independent martingales, it is easily seen that \( \hat{X}_n(t) \) and \( \hat{Z}_n(t) \) are martingales on \([0, 1]\).

**Lemma 8.1.** Suppose that \( np^d \to \infty \) and \( Ex_n \to \infty \). Then
\[ X_n - Ex_n = (1 - p^d)^{-d} \hat{X}_n(p) + 2a_2(1 - p) \hat{Z}_n(p) + \hat{R}, \]
where \( a_2 \) is defined by (5.5), and \( E\hat{R}^2 = o(\text{Var } X_n) \).

**Remark 8.1.** When \( d \geq 3 \) and \( \nu_n \to -\infty \), we may include the term with \( \hat{X}_n(p) \) in \( \hat{R} \) which gives the same result as that in Lemma 5.2. On the other hand, if \( d = 2 \) or \( \nu_n \to \infty \), then we may include the term \( 2a_2(1 - p) \hat{Z}_n(p) \) in \( \hat{R} \).

**Proof.** By cross multiplying the second product in the definition of \( \hat{X}_n(p) \), we obtain
\begin{equation}
\hat{X}_n(p) = \sum_{k=0}^{n-d} \frac{1}{d!} \frac{1}{k!} (-1)^{\frac{k}{2}} (1 - p)^{-\frac{d}{2} - dk} S(M_{d,k}), \tag{8.6}
\end{equation}
where \( S(G) \) was defined in Section 5. The argument is essentially the same as in the proof of Lemma 5.1, but much simpler, because we only have to consider terms with \( H = K_d \) and \( \nu_\beta = 0 \) for all \( \beta \neq \emptyset, \alpha \).

Now, using (8.5), (8.6) and Lemma 5.2, we can write, for some \( R \) with \( \text{Var}(R) = o(\text{Var}(X)) \),
\begin{align*}
\hat{R} &= a_2S(K_2) + \sum_{k=2}^{n-d} b_k S(M_{d,k}) + R - \sum_{k=0}^{n-d} b_k (1 - p^d)^k (1 - p)^{-\frac{d}{2} - dk} S(M_{d,k}) \\
&\quad - a_2S(K_2) \\
&= \sum_{k=2}^{n-d} \frac{(-1)^k}{d!k!} \left[(1 - p^d)^{n-d-k} - (1 - p^d)^{n-d}(1 - p)^{-\frac{d}{2} - dk}\right] S(M_{d,k}) \tag{8.7}
\end{align*}
\begin{align*}
&\quad - \sum_{k=0}^{1} b_k (1 - p^d)^k (1 - p)^{-\frac{d}{2} - dk} S(M_{d,k}) + R.
\end{align*}
The variance of the first sum on the right-hand side of (8.7) equals

\[ \begin{align*}
(1 - p^d)^{2(n-d)} & \sum_{k=2}^{n-d} \frac{1}{(d!k!)^2} \left( (1 - p^d)^{-2k} - 2(1 - p^d)^{-k} + (p^d)^{-k} \right) \\
& + (1 - p)^{-2(\xi)} \cdot d!k!(n)_{d+k} \{ p(1 - p) \}^{\xi + dk} \\
& \leq \frac{1}{d!} \binom{n}{d} p^{(\xi)} (1 - p^d)^{2(n-d)} \sum_{k=0}^{n-d} \binom{n-d}{k} (1 - p)^{(\xi)} \left( p^d \frac{1 - p^d}{1 - p^d} \right)^k - 2 \left( \frac{p^d}{1 - p^d} \right)^k \\
& + (1 - p)^{-(\xi)} \left( \frac{p^d}{1 - p^d} \right)^k \\
& = \binom{n}{d} p^{(\xi)} (1 - p^d)^{2(n-d)} \left( 1 + \frac{p^d(1 - p^d)}{(1 - p^d)^2} \right)^{n-d} - 2 \left[ 1 + \frac{p^d}{1 - p^d} \right]^{n-d} \\
& + (1 - p)^{-(\xi)} \left[ 1 + \frac{p^d}{(1 - p^d)} \right]^{n-d} \\
& = O(n^d \pi e^{-np} e^{npn} np^{d+1}) = o(n^d \pi) = o(\text{Var } X),
\end{align*} \]

by Lemma 4.1. The proof that also the variance of the last sum in (8.7) is \( o(\text{Var } X) \) is much simpler and therefore left to the reader. \( \blacksquare \)

Before proceeding, let us recall the notion of quadratic variation of a martingale [we assume for simplicity that \( M(0) = 0 \)]:

\[ [M, M] = \lim_{n \to \infty} \sum_{i=1}^{n} (M(t_i) - M(t_{i-1}))^2, \]  

(8.8)

where \( 0 = t_0 \leq t_1 \leq \cdots \leq t_n = t \), and the limit is taken in probability as \( \max(t_i - t_{i-1}) \to 0 \). (See Protter [12, Section II.6 and in particular Theorem II.22].) The martingales we use are smooth except for a finite number of jumps, and thus their quadratic variation is just the sum of the squares of the jumps:

\[ [M, M] = \sum_{s \in t} (\Delta M(s))^2, \]  

(8.9)

where \( \Delta M(s) = M(s) - M(s-) \); see Protter [12, Theorem II.26]. (The formally continuous sum in (8.9) is really a finite sum.) We recall that if \( M \) is a square integrable martingale on \([0, t]\), then \( E[M, M] = EM(t)^2 \) (Protter [12, Cor. 3 to Theorem II.27]).

Furthermore, we will also need the quadratic covariation \([M, N]\), of two martingales, which can be defined by

\[ [M, N] = \frac{1}{2} (\{M + N, M + N\} - [M, M] - [N, N]). \]  

(8.10)

In order to complete the proof of Theorem B(c) we need the following two lemmas.
Lemma 8.2. Let \((t_n)\) be a sequence of positive numbers, and assume that, for each \(n \geq 1\), \(M_n(t)\) is a martingale on \([0, t_n]\) with \(M_n(0) = 0\). If there exist deterministic continuous functions \(\sigma_n^2(t), t \in [0, t_n]\), with \(\sigma_n^2(t_n) > 0\), such that

\[
\sup_{0 \leq t \leq t_n} E[|M_n, M_n|, \sigma_n^2(t)| = o(\sigma_n^2(t_n)) \quad \text{as} \quad n \to \infty,
\]

then \(\tilde{M}_n(t_n) \to N(0, 1)\).

Lemma 8.3. Let

\[
f(t) = f_n(t) = \frac{1}{d!} n^d t^{d/2} (1 - t)^{-n}
\]

and

\[
g(t) = g_n(t) = \left(\frac{n}{2}\right) \frac{t}{1 - t}.
\]

If \(np^d \to \infty\) and \(EX_n \to \infty\), then

\[
\sup_{t \leq p} E[|\tilde{X}_n, \tilde{X}_n|, f(t)| = o(f(p))
\]

\[
\sup_{t \leq p} E[|\tilde{Z}_n, \tilde{Z}_n|, g(t)| = o(g(p))
\]

\[
\sup_{t \leq p} E[|\tilde{X}_n, \tilde{Z}_n|] = o((f(p)g(p))^{1/2}).
\]

We postpone the proofs of Lemmas 8.2 and 8.3 for a moment and first complete the proof of Theorem B.

Assume that \(np^d \to \infty\) and \(\lambda_n \to -\infty\) so that \(EX_n \to \infty\). Let \(M_n(t) = \alpha_n \tilde{X}_n(t) + \beta_n \tilde{Z}_n(t)\), where \((\alpha_n)\) and \((\beta_n)\) are any sequences of real numbers, with \(\alpha_n\) and \(\beta_n\) not both zero. Then

\[
[M_n, M_n] = \alpha_n^2 [\tilde{X}_n, \tilde{X}_n] + 2\alpha_n\beta_n [\tilde{X}_n, \tilde{Z}_n] + \beta_n^2 [\tilde{Z}_n, \tilde{Z}_n]
\]

and Lemma 8.3 yields, using the Cauchy–Schwarz inequality,

\[
\sup_{t \leq p} E[|\tilde{M}_n, \tilde{M}_n|, -\alpha_n^2 f(t) - \beta_n^2 g(t)| = o(\alpha_n^2 f(p_n) + \beta_n^2 g(p_n))
\]

\[
+ \alpha_n\beta_n (f(p_n)g(p_n))^{1/2}
\]

\[
= o(\alpha_n^2 f(p_n) + \beta_n^2 g(p_n)).
\]

Consequently, Lemma 8.2 with \(\sigma_n^2(t) = \alpha_n^2 f(t) + \beta_n^2 g(t)\) gives \(\tilde{M}_n(p_n) \to N(0, 1)\). Choosing \(\alpha_n = (1 - p^d)^{n-d}\) and \(\beta_n = 2a_2(1 - p)\), we finally obtain from this and Lemma 8.1 that \(\tilde{X}_n \to N(0, 1)\).

Proof of Lemma 8.2. We assume for simplicity that each \(\sigma_n^2(t)\) is strictly increasing with \(\sigma_n^2(0) = 0\). [This is the case in our application. The general case follows easily by replacing \(\sigma_n^2(t)\) by \(\max_{s \leq t} \sigma_n^2(s) - \sigma_n^2(0) + n^{-1}(t/t_n)\sigma_n^2(t_n)\).]
Thus $\sigma_n^2(t)/\sigma_n^2(t_n)$ is a strictly increasing map of $[0, t_n]$ onto $[0, 1]$. Let $\phi_n$ be the inverse function, and define

$$
\tilde{M}_n(t) = (\sigma_n^2(t_n))^{-1/2} M_n(\phi_n(t)), \quad 0 \leq t \leq 1.
$$

Then $\tilde{M}_n$ is a square integrable martingale on $[0, 1]$ (with respect to the appropriate $\sigma$-fields), and

$$
E[|\tilde{M}_n, \tilde{M}_n|_t - t] = (\sigma_n^2(t_n))^{-1} E[|M_n, M_n|_{\phi_n(t)} - \sigma_n^2(\phi_n(t))] = o(1),
$$

for every $t \leq 1$. Hence

$$
[\tilde{M}_n, \tilde{M}_n, \to t \quad \text{as} \quad n \to \infty, \quad 0 \leq t \leq 1,
$$

and

$$
\text{Var} \tilde{M}_n(1) = E\tilde{M}_n(1)^2 = E[\tilde{M}_n, \tilde{M}_n]_1 \to 1.
$$

Furthermore, by Doob's inequality,

$$
E(\sup_{s \leq 1} |\Delta \tilde{M}_n(s)|)^2 \leq E(2 \sup_{s \leq 1} |\tilde{M}_n(s)|)^2 \leq 16E\tilde{M}_n(1)^2 = O(1),
$$

so the variables $\sup_{s \leq 1} |\Delta \tilde{M}_n(s)|$ are uniformly integrable.

Jacod and Shiryaev [6, Theorem VIII.3.12 (applied to $\tilde{M}_n(t \wedge 1)$)] now yields $\tilde{M}_n \overset{d}{\to} W$, where $W$ is a Wiener process on $[0, 1]$. Consequently,

$$
\tilde{M}_n(t_n) = (\text{Var} \tilde{M}_n(1))^{-1/2} \tilde{M}_n(1) \overset{d}{\to} W(1) \sim N(0, 1).
$$

Proof of Lemma 8.3. We will only prove (8.13); the proofs of (8.14) and (8.15) are similar but much simpler. We begin with a formula for the quadratic variation of $\dot{X}_n$. We will use (8.9) and therefore we first compute $\Delta \dot{X}_n(t)$. We have, by (8.3),

$$
\Delta \dot{X}_n(t) = \sum_{\alpha \in [\pi]} \left( \sum_{e \in [\pi]} \Delta \dot{J}_e(t) \prod_{e' \in [\pi \setminus \{e\}]} \dot{J}_{e'}(t) \prod_{i \in \alpha} \left( 1 - \prod_{j \in \alpha} \dot{J}_{ij}(t) \right) \right)
$$

$$
- \sum_{e \in [\pi]} \dot{J}_e(t) \sum_{k \in \alpha} \sum_{l \in \alpha \cup \{k\}} \prod_{i \in \alpha} \left( 1 - \prod_{j \in \alpha \setminus \{i\}} \dot{J}_{ij}(t) \right) \prod_{j \in \alpha \setminus \{k\}} \dot{J}_{kj}(t) \Delta \dot{J}_{kl}
$$

$$
= \sum_{k < l} (\xi_{kl} - \eta_{kl} - \eta_{lk}) \Delta J_{kl},
$$

where

$$
\xi_{e}(t) = \frac{1}{1-t} \sum_{e \in \alpha \subseteq [\pi]} \sum_{e' \in [\pi \setminus \{e\}]} \dot{J}_e(t) \prod_{i \in \alpha} \left( 1 - \prod_{j \in \alpha} \dot{J}_{ij}(t) \right) = \sum_{\alpha} \xi_{e,\alpha}(t),
$$

and

$$
\eta_{kl}(t) = \frac{1}{1-t} \sum_{\alpha \subseteq [\pi]} \prod_{e \in [\pi \setminus \{k\}]} \dot{J}_e(t) \prod_{j \in \alpha \setminus \{k\}} \dot{J}_{kj}(t) \prod_{i \in \alpha \cup \{k\}} \left( 1 - \prod_{j \in \alpha \setminus \{k\}} \dot{J}_{ij}(t) \right) = \sum_{\alpha} \eta_{e,\alpha}(t).
$$
Now, since the jump $\Delta J_{kl}(s)$ is 1 when $T_{kl} = s$ and 0 otherwise, and $T_{kl}$, $1 \leq k < l \leq n$, are all distinct with probability 1,

\[
[\hat{X}_n, \hat{X}_n]_t = \sum_{s \leq t} \sum_{k < l} (\xi_{kl}(s) - \eta_{kl}(s) - \eta_{lk}(s)) \Delta J_{kl}(s) \right] \\
= \sum_{s \leq t} \sum_{k < l} (\xi_{kl}(s) - \eta_{kl}(s) - \eta_{lk}(s))^2 \Delta J_{kl}(s) \\
= \sum_{k < l} \int_0^t (\xi_{kl}(s) - \eta_{kl}(s) - \eta_{lk}(s))^2 \, dJ_{kl}(s) \\
= \sum_{k < l} \int_0^t (\xi_{kl}^2 + \eta_{kl}^2 + \eta_{lk}^2 + 2\eta_{kl} \eta_{lk} - 2\xi_{kl} \eta_{lk} - 2\xi_{kl} \eta_{lk}) \, dJ_{kl} \\
= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5 + \mathcal{L}_6.
\]

We claim that, uniformly for all $t \leq p$,

\[
E\mathcal{L}_1 = o(f(p)), \quad (8.17)
\]

\[
E\mathcal{L}_2 = E\mathcal{L}_3 = \frac{1}{2} f(t) + o(f(p)) \quad (8.18)
\]

\[
E\mathcal{L}_4 = E\mathcal{L}_5 = \frac{1}{4} f^2(t) + o(f(p)^2) \quad (8.19)
\]

\[
E\mathcal{L}_6 = o(f(p)^2). \quad (8.20)
\]

By Cauchy–Schwarz and symmetry, it then follows that

\[
E|\mathcal{L}_4| = o(f(p)),
\]

\[
E|\mathcal{L}_5| = E|\mathcal{L}_6| \leq 2E \sum_{k < l} \int_0^t |\xi_{kl} \eta_{lk}| \, dJ_{kl} \leq 2\sqrt{E\mathcal{L}_1 \cdot E\mathcal{L}_2} = o(f(p)),
\]

and

\[
E|\mathcal{L}_2 - \frac{1}{2} f(t)| = E|\mathcal{L}_4 - \frac{1}{2} f(t)| \leq (E(\mathcal{L}_3 - \frac{1}{2} f(t))^2)^{1/2} = o(f(p)),
\]

which together with (8.17) implies (8.13).

We will focus on (8.17) and (8.19), since the other two statements can be obtained in a similar way.

We start by observing that if $F(s)$ is a function of $s$ and $\{J_\varepsilon(s)\} \in [\varepsilon]$, then, for any $\varepsilon$,

\[
E \int_0^t F(s) \, dJ_\varepsilon(s) = E(F(T_\varepsilon); T_\varepsilon \leq t) = \int_0^t E(F(s)|T_\varepsilon = s) \, ds. \quad (8.21)
\]

Hence, using the fact that $\xi_\varepsilon$ is independent of $J_\varepsilon$,

\[
E\mathcal{L}_1 = \sum_\varepsilon \int_0^t E\xi_\varepsilon^2 \, ds = \sum_\varepsilon \int_0^t \sum_\alpha \sum_\beta E(\xi_{\varepsilon,\alpha}\xi_{\varepsilon,\beta}) \, ds.
\]
If $|\alpha \cap \beta| \leq d - 2$ then there exist vertices $x$ and $y$ in $\alpha - \beta$ and the variable $\hat{J}_{xy}$ will appear only once in the product $\xi_{e, \alpha} \xi_{e, \beta}$ causing $E(\xi_{e, \alpha} \xi_{e, \beta}) = 0$. If $\alpha = \beta$ and $e \in [\frac{n}{2}]$, then

\[(1-s)^2 E\xi_{e, \alpha} \xi_{e, \beta} = \prod_{e' \in [\frac{n}{2}] \setminus \{e\}} E(\hat{J}_{e'}(s))^2 \prod_{j \in \alpha} E\left(1 - \prod_{j \in \alpha} \hat{J}_{j}(s)\right)^2 \]

\[= \left(\frac{s}{1-s}\right)^{\left(\frac{\alpha}{2}\right)-1} \left(1 + \left(\frac{s}{1-s}\right)^d\right)^{n-d},\]

whereas in the case when $|\alpha \cap \beta| = d - 1$, we argue as follows. Denote by $x$ and $y$ the unique vertices in $\alpha - \beta$ and $\beta - \alpha$, respectively. Then, if $e \subset \alpha \cap \beta$,

\[(1-s)^2 E\xi_{e, \alpha} \xi_{e, \beta} = E \prod_{e' \in [\frac{n}{2}] \setminus \{e\}} (\hat{J}_{e'}(s))^2 \prod_{j \in \alpha \cap \beta} \hat{J}_{ij}(s) \hat{J}_{xy}(s) \left\{1 - \prod_{j \in \alpha} \hat{J}_{j}(s)\right\} \]

\[\times \left\{1 - \prod_{j \in \beta} \hat{J}_{j}(s)\right\} \prod_{i \in \alpha \cup \beta} \left\{1 - \prod_{j \in \alpha} \hat{J}_{ij}(s)\right\} \prod_{j \in \beta} \left\{1 - \prod_{j \in \beta} \hat{J}_{j}(s)\right\} \]

and, after cross-multiplying, the only term with nonzero expectation is that obtained by choosing the products inside the first two brackets and the 1's from all the others. Thus

\[(1-s)^2 E\xi_{e, \alpha} \xi_{e, \beta} = E \prod_{e' \in [\frac{n}{2}] \setminus \{e\}} (\hat{J}_{e'}(s))^2 = \left(\frac{s}{1-s}\right)^{\left(\frac{\alpha}{2}\right)-1}.\]

Altogether this gives

\[E \mathbb{L}_1 = \binom{n}{2} \int_0^1 \frac{1}{(1-s)^2} \left[\left(\frac{n-2}{d-2}\right) \left(\frac{s}{1-s}\right)^{\left(\frac{\alpha}{2}\right)-1} \left[1 + \left(\frac{s}{1-s}\right)^d\right]^{n-d}\right] ds\]

\[\quad + \left(\frac{n-2}{d-3, 1, 1}\right) \left(\frac{s}{1-s}\right)^{\left(\frac{\alpha}{2}\right)-1} \right] ds\]

\[= O\left(n^2 \int_0^p \left(n^{d-2}(\frac{s}{1-s})^{\left(\frac{\alpha}{2}\right)-1} + n^{d-1}(\frac{s}{1-s})^{\left(\frac{\alpha}{2}\right)-1} \right) ds\right)\]

\[= O(n^{d-1}p^{\left(\frac{\alpha}{2}\right)-d} \ e^{np^d}) = O\left(\frac{f(p)}{np^d}\right) = o(f(p)),\]

since $np^d \to \infty$, which proves (8.17). Above we made use of the following estimate:

\[\int_0^t e^{nt} ds \sim \frac{e^{nt}}{dnt^{d-1}}, \quad \text{provided} \quad nt^d \to \infty. \quad (8.22)\]

A similar argument gives

\[E \mathbb{L}_2 = \sum_{k < l} \int_0^t \sum_{\alpha} \sum_{\beta} E(\eta_{kl, \alpha} \eta_{kl, \beta}) ds,\]
where it can be shown that all terms with $\alpha \neq \beta$ vanish. Therefore we obtain

$$E \xi_2 = \sum_{k<l} \sum_{\alpha} \int_0^{t'} E \eta_{kl,\alpha}^2(s) \, ds$$

$$= \frac{n}{d!(d-1)} \int_0^{t'} \frac{1}{(1-s)^{\frac{d}{2}}} \left( s^\frac{d}{1-s} \right)^{d} \left( 1 + \left( \frac{s}{1-s} \right)^{d} \right)^{n-d-1} \, ds$$

$$= (1 + o(1)) \frac{n}{d!(d-1)} \int_0^{t'} s^{\frac{d}{1-s}} \, ds = \frac{1}{2} f(t) + o(f(p)),$$

which shows (8.18).

Now we turn to proving (8.19). By (8.21) and an analogous formula for double integrals, we have

$$E \xi_2^2 = \sum_{k<l} \sum_{k' \leq l'} E \int_0^{t'} \int_0^{t'} \eta_{kl}^2(s) \eta_{k'l'}^2(s') \, dJ_{kl}(s) \, dJ_{k'l'}(s')$$

$$= \sum_{k<l} \int_0^{t'} E \eta_{kl}^4(s) \, ds + \sum_{k<l} \sum_{k' \leq l'} \int_0^{t'} \int_0^{t'} E(\eta_{kl}^2(s) \eta_{k'l'}^2(s')) |T_{kl} = s, T_{k'l'} = s' \rangle \, ds \, ds'$$

$$= A + B.$$  

We skip the calculations for $A$ and concentrate on $B$, which can be rewritten as

$$B = \sum_{k<l} \sum_{k' \leq l'} \sum_{a_1} \sum_{a_2} \sum_{a_3} \sum_{a_4} \sum_{(k,l) \neq (k',l')} \int_0^{t'} \int_0^{t'} E(\eta_{kl,a_1}(s) \eta_{kl,a_2}(s) \eta_{k'l',a_3}(s') \eta_{k'l',a_4}(s') | T_{kl} = s, T_{k'l'} = s' \rangle \, ds \, ds'.  \quad (8.24)$$

We are going to show that only those terms where $\alpha_1 = \alpha_2, \alpha_3 = \alpha_4$, and $\alpha_1, \alpha_3, \{k\}, \{k'\}$ are disjoint, give a significant contribution, say $B_1$, whereas the others are together $o(f(p)^2)$. In the former case the variables $\eta_{kl,a_1}$, etc. are independent of $J_{kl}$ and $J_{k'l'}$, so the conditional expectation above is just the expectation, and each term can be written as

$$\int_0^{t'} \int_0^{t'} E(\eta_{kl,a_1}^2(s) \eta_{k'l',a_3}^2(s')) \, ds \, ds'.$$

The only terms in $\eta_{kl,a_1}^2$ and $\eta_{k'l',a_3}^2$ which cause dependence are those containing $J_e$ for $e$ with one endpoint in $\alpha_1$ and the other in $\alpha_3$. We thus express them as

$$\eta_{kl,a_1} = \tilde{\eta}_{kl,a_1} \prod_{i \in a_1} \left( 1 - \prod_{j \in a_1} \tilde{J}_{ij} \right) = \tilde{\eta}_{k'l',a_3} \xi(\alpha_3, \alpha_1)$$

and

$$\eta_{k'l',a_3} = \tilde{\eta}_{k'l',a_3} \prod_{i \in a_3} \left( 1 - \prod_{j \in a_3} \tilde{J}_{ij} \right) = \tilde{\eta}_{k'l',a_3} \xi(\alpha_1, \alpha_3).$$
Since $E \xi^2(\alpha_1, \alpha_1) = 1 + O(p)$ and $E(\xi^2(\alpha_1, \alpha_1)\xi^2(\alpha_3, \alpha_3)) = 1 + O(p)$, and $\eta_{kl, \alpha_1}, \eta_{kl, \alpha_3}$, $(\xi(\alpha_1, \alpha_3), \xi(\alpha_3, \alpha_1))$ are mutually independent, we obtain

$$E\eta^2_{kl, \alpha_1}(s)\eta^2_{kl, \alpha_3}(s') = E\eta_{kl, \alpha_1}(s)E\eta^2_{kl, \alpha_3}(s')(1 + O(p)).$$

Comparing this to (8.23) we see that

$$B_1 = (E\xi^2)^2(1 + O(p)) + O\left(\frac{1}{n}\right).$$

(8.25)

It remains to show that the sum of the remaining terms in (8.24) is small. We consider one of these terms and let $V = \alpha_1 \cup \alpha_2 \cup \alpha_3 \cup \alpha_4 \cup \{k', k''\}$, and cross-multiply all factors in the $\eta$'s of the form $(1 - \Pi_{j \in \gamma} \hat{J}(s))$ with $i \in V$. This gives at most $2^{4d+2}$ terms of the type

$$\pm \int_0^t \int_0^t E\left(\prod_{\epsilon \in \{s, s'\}} \hat{J}\xi, \xi^2(s, s')^m \prod_{\epsilon \in \gamma} \xi_\epsilon \mid T_{kl} = s, T_{k,l'} = s'\right) ds \, ds', \tag{8.26}$$

where $m_\epsilon, m'_\epsilon$ are 0, 1 or 2 and, if we write $s_1 = s_2 = s, s_3 = s_4 = s'$,

$$\xi = \prod_{\epsilon = 1}^4 \left(1 - \prod_{j \in \gamma} \hat{J}(s_\epsilon)\right).$$

Let us study the conditional expectation in (8.26) further. The conditioning only affects the factors with $\epsilon = \{k, l\}$ or $\{k', l'\}$, and these two factors are $O(1)$. For every other $\epsilon$, it follows easily from the definition of $\hat{J}$ that, assuming as we may that $s, s' \leq t \leq p \leq 1/2,$

$$E\hat{J}_\epsilon(s)\xi, \xi^2(s')^m = \begin{cases} 
1 & \text{if } m_\epsilon = m'_\epsilon = 0 \\
0 & \text{if } m_\epsilon + m'_\epsilon = 1 \\
\pm s + O(s^2) & \text{if } m_\epsilon = 2, m'_\epsilon = 0 \\
\pm s' + O(s'^2) & \text{if } m_\epsilon = 0, m'_\epsilon = 2 \\
\pm (s \land s' + O(ss')) & \text{if } m_\epsilon \geq 1, m'_\epsilon \geq 1. 
\end{cases} \tag{8.27}$$

Furthermore, all factors are independent, and the $\xi_i$ all have the same distribution.

Consequently, if $\mathcal{E} = \{\epsilon \in \{1\} : m_\epsilon + m'_\epsilon > 0, \epsilon \neq \{k, l\} \text{ and } \epsilon \neq \{k', l'\}\}$, and $\epsilon = |\mathcal{E}|$, then the integrand in (8.26) is $O(t^\epsilon (E\xi)^{\gamma - |\mathcal{E}|})$ (for any $i \not\in \mathcal{E}$), and it vanishes if $m_\epsilon + m'_\epsilon = 1$ for some $\epsilon \in \mathcal{E}$.

We now consider the case $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$ separately, and let $B_2$ denote the total contribution of such terms to (8.24), while $B_3$ denotes the contribution from the remaining terms. (Thus $B = B_1 + B_2 + B_3$.)

If $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$ we obtain, by cross-multiplying in the definition of $\alpha_i$ and using (8.27),

$$E\xi = 1 - 4 \cdot 0 + s^d + 4(s \land s')^d + s'^d - 4(s \land s')^d + (s \land s')^d + O(t^{d+1})$$

$$= 1 + s^d + s'^d + (s \land s')^d + O(t^{d+1}) \leq 1 + s^d + s'^d + t^d + O(t^{d+1}),$$

(8.28)
and thus, since \( nt^{d+1} \leq np^{d+1} \to 0 \),
\[
(E \xi_i)^{n-|V|} = O(e^{nt^{d+ns^{d}+nt^{d}}}).
\]
Consequently, each term (8.26) is, using also (8.22),
\[
O\left(\int_0^r \int_0^r t^i e^{nt^{d+ns^{d}+nt^{d}}} dt \, ds\right) = O\left(p^n e^{npd} \left(\int_0^p e^{nt^{d}} dt\right)^2\right)
\]
\[
= O(n^{-2} p^{-2(d-1)} e^{3npd}).
\]
Now there are \( O(n^{d+2}) \) such terms with \( k \neq k' \), each having \( e = \left(\frac{d}{2}\right) + 2(d-1) \), and \( O(n^{d+1}) \) such terms with \( k = k' \) and \( e \geq \left(\frac{d}{2}\right) + d - 2 \). This yields
\[
B_2 = O((1 + (np^{d})^{-1})n^d p^{(\frac{d}{2})} e^{3npd}) = O(n^d p^{(\frac{d}{2})} e^{3npd}) = O\left(\frac{1}{EX} f(p)^2\right)
\]
\[
= o(f(p)^2), \quad (8.29)
\]
because \( EX = n^d p^{(\frac{d}{2})} e^{-npd} \to \infty \).

Finally assume that not all \( \alpha_i \) are equal. \( E \xi_i \) is easily computed up to \( O(t^{d+1}) \) as above. The exact result depends on which, if any, of \( \alpha_1, \ldots, \alpha_4 \) that coincide, but in all cases
\[
E \xi_i \leq 1 + s^d + s^{rd} + O(t^{d+1}) \quad (8.30)
\]
and thus each term (8.26) is
\[
O\left(\int_0^r \int_0^r t^i e^{nt^{d+ns^{d}}} dt \, ds\right) = O(n^{-2} p^{-2(d-1)} e^{2npd})
\]
\[
= O(n^{-2-2d} p^{e+2-2d-2(\frac{d}{2})} f(p)^2). \quad (8.31)
\]
We now claim that for any nonvanishing term (8.26)
\[
v = |V| \leq 2d + 2, \quad (8.32)
\]
\[
e + 2 \geq dv/2, \quad (8.33)
\]
and that strict inequality holds in at least one of (8.32) and (8.33) unless we are in a case included in the sum \( B_1 \). In order to show (8.32), assume first that \( k \not\in \alpha_3 \cup \alpha_4 \). If \( \alpha_1 \) or \( \alpha_2 \) contains a vertex, \( x \) say, not contained in any other \( \alpha_i \), then \( \vec{J}_k \) appears exactly once in the product in (8.26), and the expectation vanishes. If also \( k' \not\in \alpha_3 \cup \alpha_2 \), the same argument shows that every vertex in any \( \alpha_v \) belongs to at least two of them, and thus
\[
v \leq 2 + \left| \bigcup_{\nu=1}^{4} \alpha_\nu \right| \leq 2 + \frac{1}{2} \sum_{\nu=1}^{4} |\alpha_\nu| = 2 + 2d.
\]
On the other hand, if \( k' \in \alpha_1 \cup \alpha_2, \alpha_3 \) and \( \alpha_4 \) may contain at most one vertex each
not in any other $\alpha_v$, because if, say, $x, y \in \alpha_1 \cup \ldots \cup \alpha_4$, then $J_{xy}$ appears exactly once in the product. Consequently,

$$v \leq 1 + \left| \bigcup_{v=1}^{4} \alpha_v \right| \leq 1 + 2 + \frac{1}{2} (4d - 2) = 2d + 2.$$  

The case $k \in \alpha_3 \cup \alpha_4$ is similar.

For (8.33), let us regard $\mathcal{E} \cup \{k, l\} \cup \{k', l'\}$ as the edges of a graph with $\varepsilon + 2$ edges and $v$ vertices; then this graph is a union of four $K_{d+1}$'s, with possibly further edges added. Consequently, each vertex has degree at least $d$, and thus $2(\varepsilon + 2) \geq vd$, with equality only if the four sets $\alpha_v \cup \{k_v\}$ (where $k_1 = k_2 = k$, $k_3 = k_4 = k'$) pairwise either coincide or are disjoint. This proves (8.33), with equality only when $\alpha_1 = \alpha_2$, $\alpha_3 = \alpha_4$ and $\alpha_v \cup \{k\}$ is disjoint from $\alpha_v \cup \{k'\}$ (and this term is included in $B_1$), or $\alpha_v \cup \{k_v\}$ are all equal and $v = d + 1 < 2d + 2$.

Finally, since there are $O(n^v)$ terms with a given $v$, their total contribution to $B_3$ is by (8.31)–(8.33), if $v < 2d + 2$,

$$O(n^{v-2-2d} p^{dv/2-d(d+1)} f(p)^2) = O((np)^{d-(2d+2-v)} f(p)^2) = o(f(p)^2)$$

and, if $v = 2d + 2$, when strict inequality holds in (8.33),

$$O(p^{\varepsilon+2-dv/2} f(p)^2) = o(f(p)^2).$$

Consequently, $B_3 = o(f(p)^2)$, which together with (8.25) and (8.29) yield (8.19).

A similar argument shows (8.20).

\section*{Acknowledgment}

This work was partially supported under grant 21-25579.88 by the Swiss Nationalfonds whose assistance is gratefully acknowledged.

\section*{References}


Received March 1990