Proceedings of the 27th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms
Kraków, Poland, 4-8 July 2016

On the Higher Dimensional Quasi-Power Theorem and a Berry–Esseen Inequality

Clemens Heuberger[1] [†] and Sara Kropf[1,2] [†]

[1] *Institut für Mathematik, Alpen-Adria-Universität Klagenfurt, Austria,* {*clemens.heuberger, sara.kropf*}*@aau.at*
[2] *Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, sarakropf@stat.sinica.edu.tw*

Hwang's quasi-power theorem asserts that a sequence of random variables whose moment generating functions are approximately given by powers of some analytic function is asymptotically normally distributed. This theorem is generalised to higher dimensional random variables. To obtain this result, a higher dimensional analogue of the Berry–Esseen inequality is proved, generalising a two-dimensional version by Sadikova.

Keywords: Quasi-power theorem, Berry–Esseen inequality, limiting distribution, central limit theorem

1 Introduction

Asymptotic normality is a frequently occurring phenomenon in combinatorics, the classical central limit theorem being the very first example. The first step in the proof is the observation that the moment generating function of the sum of n identically independently distributed random variables is the n-th power of the moment generating function of the distribution underlying the summands. As similar moment generating functions occur in many examples in combinatorics, a general theorem to prove asymptotic normality is desirable. Such a theorem was proved by Hwang [16], usually called the "quasi-power theorem".

Theorem (Hwang [16])**.** *Let $\{\Omega_n\}_{n\geq 1}$ be a sequence of integral random variables. Suppose that the moment generating function satisfies the asymptotic expression*

$$M_n(s) := \mathbb{E}(e^{\Omega_n s}) = e^{W_n(s)}(1 + O(\kappa_n^{-1})), \tag{1.1}$$

the O-term being uniform for $|s| \leq \tau$, $s \in \mathbb{C}$, $\tau > 0$, where

1. *$W_n(s) = u(s)\phi_n + v(s)$, with $u(s)$ and $v(s)$ analytic for $|s| \leq \tau$ and independent of n; and $u''(0) \neq 0$;*
2. *$\lim_{n\to\infty} \phi_n = \infty$;*
3. *$\lim_{n\to\infty} \kappa_n = \infty$.*

[†] The authors are supported by the Austrian Science Fund (FWF): P 24644-N26.

Then the distribution of Ω_n is asymptotically normal, i.e.,

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(\frac{\Omega_n - u'(0)\phi_n}{\sqrt{u''(0)\phi_n}} < x\right) - \Phi(x)\right| = O\left(\frac{1}{\sqrt{\phi_n}} + \frac{1}{\kappa_n}\right),$$

where Φ denotes the standard normal distribution

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-\frac{1}{2}y^2\right)\,dy.$$

See Hwang's article [16] as well as Flajolet-Sedgewick [8, Sec. IX.5] for many applications of this theorem. A generalisation of the quasi-power theorem to dimension 2 has been provided in [12]. It has been used in [14], [15], [6], [13] and [17]. In [5, Thm. 2.22], an m-dimensional version of the quasi-power theorem is stated without speed of convergence. Also in [2], such an m-dimensional theorem without speed of convergence is proved. There, several multidimensional applications are given, too.

In contrast to many results about the speed of convergence in classical probability theory (see, e.g., [11]), the sequence of random variables is not assumed to be independent. The only assumption is that the moment generating function behaves asymptotically like a large power. This mirrors the fact that the moment generating function of the sum of independent, identically distributed random variables is exactly a large power. The advantage is that the asymptotic expression (1.1) arises naturally in combinatorics by using techniques such as singularity analysis or saddle point approximation (see [8]).

The purpose of this article is to generalise the quasi-power theorem including the speed of convergence to arbitrary dimension m. We first state this main result in Theorem 1 in this section. In Section 2, a new Berry–Esseen inequality (Theorem 2) is presented, which we use to prove the m-dimensional quasi-power theorem. We give sketches of the proofs of these two theorems in Section 4. All details of these proofs can be found in the full version of this extended abstract. In Section 3, we give some applications of the multidimensional quasi-power theorem.

We use the following conventions: vectors are denoted by boldface letters such as $\mathbf{s}$, their components are then denoted by regular letters with indices such as s_j. For a vector $\mathbf{s}$, $\|\mathbf{s}\|$ denotes the maximum norm $\max\{|s_j|\}$. All implicit constants of O-terms may depend on the dimension m as well as on τ which is introduced in Theorem 1.

Our first main result is the following m-dimensional version of Hwang's theorem.

Theorem 1. *Let $\{\mathbf{\Omega}_n\}_{n\geq 1}$ be a sequence of m-dimensional real random vectors. Suppose that the moment generating function satisfies the asymptotic expression*

$$M_n(\mathbf{s}) := \mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s}\rangle}) = e^{W_n(\mathbf{s})}(1 + O(\kappa_n^{-1})), \tag{1.2}$$

the O-term being uniform for $\|\mathbf{s}\| \leq \tau$, $\mathbf{s} \in \mathbb{C}^m$, $\tau > 0$, where

1. *$W_n(\mathbf{s}) = u(\mathbf{s})\phi_n + v(\mathbf{s})$, with $u(\mathbf{s})$ and $v(\mathbf{s})$ analytic for $\|\mathbf{s}\| \leq \tau$ and independent of n; and the Hessian $H_u(\mathbf{0})$ of u at the origin is non-singular;*

2. *$\lim_{n\to\infty} \phi_n = \infty$;*

3. *$\lim_{n\to\infty} \kappa_n = \infty$.*

Then, the distribution of $\mathbf{\Omega}_n$ is asymptotically normal with speed of convergence $O({\phi_n}^{-1/2})$, i.e.,

$$\sup_{\mathbf{x}\in\mathbb{R}^m}\left|\mathbb{P}\left(\frac{\mathbf{\Omega}_n-\operatorname{grad}u(\mathbf{0})\phi_n}{\sqrt{\phi_n}}\le\mathbf{x}\right)-\Phi_{H_u(\mathbf{0})}(\mathbf{x})\right|=O\left(\frac{1}{\sqrt{\phi_n}}\right), \tag{1.3}$$

where Φ_Σ denotes the distribution function of the non-degenerate m-dimensional normal distribution with mean $\mathbf{0}$ and variance-covariance matrix Σ, i.e.,

$$\Phi_\Sigma(\mathbf{x})=\frac{1}{(2\pi)^{m/2}\sqrt{\det\Sigma}}\int_{\mathbf{y}\le\mathbf{x}}\exp\left(-\frac{1}{2}\mathbf{y}^\top\Sigma^{-1}\mathbf{y}\right)\,d\mathbf{y},$$

where $\mathbf{y}\le\mathbf{x}$ means $y_\ell\le x_\ell$ for $1\le\ell\le m$.

If $H_u(\mathbf{0})$ is singular, the random variables

$$\frac{\mathbf{\Omega}_n-\operatorname{grad}u(\mathbf{0})\phi_n}{\sqrt{\phi_n}}$$

converge in distribution to a degenerate normal distribution with mean $\mathbf{0}$ and variance-covariance matrix $H_u(\mathbf{0})$.

Note that in the case of the singular $H_u(\mathbf{0})$, a uniform speed of convergence cannot be guaranteed. To see this, consider the (constant) sequence of random variables Ω_n which takes values ±1 each with probability $1/2$. Then the moment generating function is $(e^t+e^{-t})/2$, which is of the form (1.2) with $\phi_n=n$, $u(s)=0$, $v(s)=\log(e^t+e^{-t})/2$ and κ_n arbitrary. However, the distribution function of $\Omega_n/\sqrt{n}$ is given by

$$\mathbb{P}\left(\frac{\Omega_n}{\sqrt{n}}\le x\right)=\begin{cases}0 & \text{if } x<-1/\sqrt{n},\\ 1/2 & \text{if } -1/\sqrt{n}\le x<1/\sqrt{n},\\ 1 & \text{if } 1/\sqrt{n}\le x,\end{cases}$$

which does not converge uniformly.

In contrast to the original quasi-power theorem, the error term in our result does not contain the summand $O(1/\kappa_n)$. In fact, this summand could also be omitted in the original proof of the quasi-power theorem by using a better estimate for the error $E_n(\mathbf{s})=M_n(\mathbf{s})e^{-W_n(\mathbf{s})}-1$.

The proof of Theorem 1 relies on an m-dimensional Berry–Esseen inequality (Theorem 2). It is a generalisation of Sadikova's result [22, 23] in dimension 2. The main challenge is to provide a version which leads to bounded integrands around the origin, but still allows to use excellent bounds for the tails of the characteristic functions. To achieve this, linear combinations involving all partitions of the set $\{1,\ldots,m\}$ are used.

Note that there are several generalisations of the one-dimensional Berry–Esseen inequality [3, 7] to arbitrary dimension, see, e.g., Gamkrelidze [9, 10] and Prakasa Rao [20]. However, using these results would lead to the less precise error term in (1.3), see the end of Section 2 for more details. For that reason we generalise Sadikova's result, which was already successfully used by the first author in [12] to prove a 2-dimensional quasi-power theorem. Also note that our theorem can deal with discrete random variables, in contrast to [21], where density functions are considered.

For the sake of completeness, we also state the following result about the moments of $\mathbf{\Omega}_n$.

Proposition 1.1. *The cross-moments of $\mathbf{\Omega}_n$ satisfy*

$$\frac{1}{\prod_{\ell=1}^{m} k_\ell!}\mathbb{E}\Big(\prod_{\ell=1}^{m}\Omega_{n,\ell}^{k_\ell}\Big) = p_{\mathbf{k}}(\phi_n) + O\Big(\kappa_n^{-1}\phi_n^{k_1+\cdots+k_m}\Big),$$

for k_ℓ nonnegative integers, where $p_{\mathbf{k}}$ is a polynomial of degree $\sum_{\ell=1}^{m} k_\ell$ defined by

$$p_{\mathbf{k}}(X) = [s_1^{k_1}\cdots s_m^{k_m}]e^{u(\mathbf{s})X+v(\mathbf{s})}.$$

In particular, the mean and the variance-covariance matrix are

$$\begin{aligned}\mathbb{E}(\mathbf{\Omega}_n) &= \operatorname{grad} u(\mathbf{0})\phi_n + \operatorname{grad} v(\mathbf{0}) + O(\kappa_n^{-1}),\\ \operatorname{Cov}(\mathbf{\Omega}_n) &= H_u(\mathbf{0})\phi_n + H_v(\mathbf{0}) + O(\kappa_n^{-1}),\end{aligned}$$

respectively.

2 A Berry–Esseen Inequality

This section is devoted to a generalisation of Sadikova's Berry–Esseen inequality [22, 23] in dimension 2 to dimension m. Before stating the theorem, we introduce our notation.

Let $L = \{1, \dots, m\}$. For $K \subseteq L$, we write $\mathbf{s}_K = (s_k)_{k\in K}$ for the projection of $\mathbf{s} \in \mathbb{C}^L$ to $\mathbb{C}^K$. For $J \subseteq K \subseteq L$, let $\chi_{J,K} : \mathbb{C}^J \to \mathbb{C}^K$, $(s_j)_{j\in J} \mapsto (s_k[k \in J])_{k\in K}$ be an injection from $\mathbb{C}^J$ into $\mathbb{C}^K$. Similarly, let $\psi_{J,K} : \mathbb{C}^K \to \mathbb{C}^K$, $(s_k)_{k\in K} \mapsto (s_k[k \in J])_{k\in K}$ be the projection which sets all coordinates corresponding to $K \setminus J$ to 0.

We denote the set of all partitions of K by Π_K. We consider a partition as a set $\alpha = \{J_1, \dots, J_k\}$. Thus $|\alpha|$ denotes the number of parts of the partition α. Furthermore, $J \in \alpha$ means that J is a part of the partition α.

Now, we can define an operator which we later use to state our Berry–Esseen inequality. The motivation behind this definition is explained at the end of this section.

Definition 2.1. Let $K \subseteq L$ and $h : \mathbb{C}^K \to \mathbb{C}$. We define the non-linear operator

$$\Lambda_K(h) := \sum_{\alpha\in\Pi_K} \mu_\alpha \prod_{J\in\alpha} h \circ \psi_{J,K}$$

where

$$\mu_\alpha = (-1)^{|\alpha|-1}(|\alpha| - 1)!\,.$$

We denote Λ_L briefly by Λ.

For any random variable $\mathbf{Z}$, we denote its cumulative distribution function by $F_{\mathbf{Z}}$ and its characteristic function by $\varphi_{\mathbf{Z}}$.

With these definitions, we are able to state our second main result, an m-dimensional version of the Berry–Esseen inequality.

Theorem 2. *Let $m \geq 1$ and $\mathbf{X}$ and $\mathbf{Y}$ be m-dimensional random variables. Assume that $F_{\mathbf{Y}}$ is differentiable.*

Let

$$\begin{aligned} A_j &= \sup_{\mathbf{y}\in\mathbb{R}^m} \frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j}, \\ B_j &= \sum_{k=1}^{j} \left\{ {j \atop k} \right\} k!\,, \\ C_1 &= \sqrt[3]{\frac{32}{\pi\bigl(1-\bigl(\frac{3}{4}\bigr)^{1/m}\bigr)}}, \\ C_2 &= \frac{12}{\pi} \end{aligned}$$

for $1 \le j \le m$ *where* $\left\{ {j \atop k} \right\}$ *denotes a Stirling partition number (Stirling number of the second kind).*

Let $T > 0$ *be fixed. Then*

$$\begin{aligned} \sup_{\mathbf{z}\in\mathbb{R}^m} |F_{\mathbf{X}}(\mathbf{z}) - F_{\mathbf{Y}}(\mathbf{z})| \le \frac{2}{(2\pi)^m} &\int_{\|\mathbf{t}\|\le T} \left| \frac{\Lambda(\varphi_{\mathbf{X}})(\mathbf{t}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{t})}{\prod_{\ell\in L} t_\ell} \right| d\mathbf{t} \\ &+ 2 \sum_{\emptyset\ne J\subsetneq L} B_{m-|J|} \sup_{\mathbf{z}_J\in\mathbb{R}^J} \bigl|F_{\mathbf{X}_J}(\mathbf{z}_J) - F_{\mathbf{Y}_J}(\mathbf{z}_J)\bigr| \\ &+ \frac{2\sum_{j=1}^{m} A_j}{T}(C_1 + C_2). \end{aligned} \tag{2.1}$$

Existence of $\mathbb{E}(\mathbf{X})$ *and* $\mathbb{E}(\mathbf{Y})$ *is sufficient for the finiteness of the integral in* (2.1).

Let us give two remarks on the distribution functions occurring in this theorem: The distribution function $F_{\mathbf{Y}}$ is non-decreasing in every variable, thus $A_j > 0$ for all j. Furthermore, our general notations imply that $F_{\mathbf{X}_J}$ is a marginal distribution of $\mathbf{X}$.

The numbers B_j are known as "Fubini numbers" or "ordered Bell numbers". They form the sequence A000670 in [18].

Recursive application of (2.1) leads to the following corollary, where we no longer explicitly state the constants depending on the dimension.

Corollary 2.2. *Let* $m \ge 1$ *and* $\mathbf{X}$ *and* $\mathbf{Y}$ *be* m*-dimensional random variables. Assume that* $F_{\mathbf{Y}}$ *is differentiable and let*

$$A_j = \sup_{\mathbf{y}\in\mathbb{R}^m} \frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j}, \qquad 1 \le j \le m.$$

Then

$$\begin{aligned} &\sup_{\mathbf{z}\in\mathbb{R}^m} |F_{\mathbf{X}}(\mathbf{z}) - F_{\mathbf{Y}}(\mathbf{z})| \\ &\quad = O\biggl(\sum_{\emptyset\ne K\subseteq L} \int_{\|\mathbf{t}_K\|\le T} \left| \frac{\Lambda_K(\varphi_{\mathbf{X}}\circ\chi_{K,L})(\mathbf{t}_K) - \Lambda_K(\varphi_{\mathbf{Y}}\circ\chi_{K,L})(\mathbf{t}_K)}{\prod_{k\in K} t_k} \right| d\mathbf{t}_K + \frac{\sum_{j=1}^{m} A_j}{T} \biggr) \end{aligned} \tag{2.2}$$

where the O*-constants only depend on the dimension* m.

Existence of $\mathbb{E}(\mathbf{X})$ *and* $\mathbb{E}(\mathbf{Y})$ *is sufficient for the finiteness of the integrals in* (2.2).

In order to explain the choice of the operator Λ, we first state it in dimension 2:

$$\Lambda(h)(s_1, s_2) = h(s_1, s_2) - h(s_1, 0)h(0, s_2). \tag{2.3}$$

This coincides with Sadikova's definition. This also shows that our operator is non-linear as, e.g., $\Lambda(s_1 + s_2)(s_1, s_2) \neq \Lambda(s_1)(s_1, s_2) + \Lambda(s_2)(s_1, s_2)$.

In Theorem 2, we apply Λ to characteristic functions; so we may restrict our attention to functions h with $h(\mathbf{0}) = 1$. From (2.3), we see that $\Lambda(h)(s_1, 0) = \Lambda(h)(0, s_2) = 0$, so that $\Lambda(h)(s_1, s_2)/(s_1 s_2)$ is bounded around the origin. This is essential for the boundedness of the integral in Theorem 2. In general, this property will be guaranteed by our particular choice of coefficients. It is no coincidence that for $\alpha \in \Pi_L$, the coefficient μ_α equals the value $\mu(\alpha, \{L\})$ of the Möbius function in the lattice of partitions: Weisner's theorem (see Stanley [24, Corollary 3.9.3]) is crucial in the proof that $\Lambda(h)(\mathbf{s})/(s_1 \cdots s_m)$ is bounded around the origin.

The second property is that our proof of the quasi-power theorem needs estimates for the tails of the integral in Theorem 2. These estimates have to be exponentially small in every variable, which means that every variable has to occur in every summand. This is trivially fulfilled as every summand in the definition of Λ is formulated in terms of a partition.

Note that Gamkrelidze [10] (and also Prakasa Rao [20]) use a linear operator L mapping h to

$$(s_1, s_2) \mapsto h(s_1, s_2) - h(s_1, 0) - h(0, s_2). \tag{2.4}$$

When taking the difference of two characteristic functions, we may assume that $h(0, 0) = 0$ so that the first crucial property as defined above still holds. However, the tails are no longer exponentially small in every variable: The last summand $h(0, s_2)$ in (2.4) is not exponentially small in s_1 because it is independent of s_1 and nonzero in general. However, the first two summands are exponentially small in s_1 by our assumption (1.2).

For that reason, using the Berry–Esseen inequality by Gamkrelidze [10] to prove a quasi-power theorem leads to a less precise error term $O(\phi_n^{-1/2} \log^{m-1} \phi_n)$ in (1.3). It can be shown that the less precise error term necessarily appears when using Gamkrelidze's result by considering the example of $\mathbf{\Omega}_n$ being the 2-dimensional vector consisting of a normal distribution with mean -1 and variance n and a normal distribution with mean 0 and variance n. This is a consequence of the linearity of the operator L in Gamkrelidze's result.

3 Examples of Multidimensional Central Limit Theorems

In this section, we give two examples from combinatorics where we can apply Theorem 1. Asymptotic normality was already shown in earlier publications [4, 2], but we additionally provide an estimate for the speed of convergence.

3.1 Context-Free Languages

Consider the following example of a context-free grammar G with non-terminal symbols S and T, terminal symbols $\{a, b, c\}$, starting symbol S and the rules

$$P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}.$$

The corresponding context-free language $L(G)$ consists of all words which can be generated starting with S using the rules in P to replace all non-terminal symbols. For example, $abcabababba \in L(G)$ because it can be derived as

$$S \to aSbS \to abTbaSbS \to abcTbabTbbT \to abcabababba.$$

Let $\mathbb{P}(\mathbf{\Omega}_n = \mathbf{x})$ be the probability that a word of length n in $L(G)$ consists of x_1 and x_2 terminal symbols a and b, respectively. Thus there are $n - x_1 - x_2$ terminal symbols c. For simplicity, this random variable is only 2-dimensional. But it can be easily extended to higher dimensions.

Following Drmota [4, Sec. 3.2], we obtain that the moment generating function is

$$\mathbb{E}(e^{\langle \mathbf{\Omega}_n, \mathbf{s} \rangle}) = \frac{y_n(e^{\mathbf{s}})}{y_n(\mathbf{1})}$$

with $y_n(\boldsymbol{z})$ defined in [4]. Using [4, Equ. (4.9)], this moment generating function has an asymptotic expansion as in (1.2) with $\phi_n = n$. Thus $\mathbf{\Omega}_n$ is asymptotically normally distributed after standardisation (as was shown in [4]) and additionally the speed of convergence is $O(n^{-1/2})$.

Other context-free languages can be analysed in the same way, either by directly using the results in [4] (if the underlying system is strongly connected) or by similar methods. This has applications, for example, in genetics (see [19]).

3.2 Dissections of Labelled Convex Polygons

Let $S_1 \uplus \cdots \uplus S_{t+1} = \{3, 4, \ldots\}$ be a partition. We dissect a labelled convex n-gon into smaller convex polygons by choosing some non-intersecting diagonals. Each small polygon should be a k-gon with $k \notin S_{t+1}$. Define $a_n(\mathbf{r})$ to be the number of dissections of an n-gon such that it consists of exactly r_i small polygons whose number of vertices is in S_i, for $i = 1, \ldots, t$. For convenience, we use $a_2(\mathbf{r}) = [\mathbf{r} = \mathbf{0}]$. Asymptotic normality was proved in [2, Sec. 3], see also [1, Ex. 7.1] for a one-dimensional version. We additionally provide an estimate for the speed of convergence.

Let

$$f(z, \mathbf{x}) = \sum_{\substack{n \geq 2 \\ \mathbf{r} \geq 0}} a_n(\mathbf{r}) \mathbf{x}^{\mathbf{r}} z^{n-1}.$$

Then choosing a k-gon with $k \in S_1 \uplus \cdots \uplus S_t$ and gluing dissected polygons to $k-1$ of its sides translates into the equation

$$f = z + \sum_{i=1}^{t} x_i \sum_{k \in S_i} f^{k-1}.$$

Following [1], this equation can be used to obtain an asymptotic expression for the moment generating function as in (1.2) with $\phi_n = n$. The asymptotic normal distribution follows after suitable standardisation with speed of convergence $O(n^{-1/2})$.

4 Sketch of the Proofs

We now sketch the main ideas of the proofs of Theorems 2 and 1. All details can be found in the full version of this extended abstract.

Sketch of the proof of Theorem 2: As in [23, 10, 20], our proof of the Berry–Esseen inequality proceeds via adding a continuous random variable $\mathbf{Q}$ to our random variables $\mathbf{X}$ and $\mathbf{Y}$. The characteristic function of $\mathbf{Q}$ vanishes outside $[-T,T]^m$. The error resulting from replacing the difference of the distribution functions $|F_{\mathbf{X}} - F_{\mathbf{Y}}|$ by $|F_{\mathbf{X}+\mathbf{Q}} - F_{\mathbf{Y}+\mathbf{Q}}|$ can be estimated by the final summand in (2.1). In principle, Lévy's theorem then allows to bound the difference of the distribution functions by the difference of the characteristic functions. Instead of only using the difference of the characteristic functions, we use the difference $|\Lambda(\varphi_{\mathbf{X}}) - \Lambda(\varphi_{\mathbf{Y}})|$, which ensures boundedness of the integral in (2.1) at least if the first moments exist. However, we have to compensate Λ by the sum over the differences of the marginal distribution functions, which yields the second summand in (2.1). □

Sketch of the proof of Theorem 1: First, the characteristic function of the standardised random variable $\mathbf{X} = (\mathbf{\Omega}_n - \operatorname{grad} u(\mathbf{0})\phi_n)/\sqrt{\phi_n}$ is

$$\varphi_{\mathbf{X}}(\mathbf{s}) = \exp\Big(-\frac{1}{2}\mathbf{s}^\top\Sigma\mathbf{s} + O\Big(\frac{\|\mathbf{s}\|^3 + \|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)\Big)$$

for $\|\mathbf{s}\| < \tau\sqrt{\phi_n}/2$. Thus, we obtain convergence in distribution as stated in the theorem.

To obtain a bound for the speed of convergence, we use the Berry–Esseen inequality given in Theorem 2 for $\mathbf{Y}$ an m-dimensional normal distribution. We bound the difference of Λ evaluated at the characteristic function of $\mathbf{X}$ and the one of the normal distribution by the exponentially decreasing function

$$|\Lambda(\varphi_{\mathbf{X}})(\mathbf{s}) - \Lambda(\varphi_{\mathbf{Y}})(\mathbf{s})| \le \exp\Big(-\frac{\sigma}{4}\|\mathbf{s}\|^2 + O(\|\mathbf{s}\|)\Big)O\Big(\frac{\|\mathbf{s}\|^3 + \|\mathbf{s}\|}{\sqrt{\phi_n}}\Big)$$

for suitable $\mathbf{s}$ where σ is the smallest eigenvalue of Σ.

We then estimate the integral in (2.1). For the variables in a neighbourhood of zero, we get rid of the denominator by Taylor expansion using the zero of $\Lambda(\varphi_{\mathbf{X}}) - \Lambda(\varphi_{\mathbf{Y}})$ at $\mathbf{0}$. The error term of the Taylor expansion can be estimated by the difference of the characteristic functions using Cauchy's formula. The exponentially small tails are used to bound the contribution of the large variables in the integral in (2.1).

The second summand in (2.1) can be estimated inductively. □

References

[1] Edward A. Bender, *Asymptotic methods in enumeration*, SIAM Rev. **16** (1974), no. 4, 485–515.

[2] Edward A. Bender and L. Bruce Richmond, *Central and local limit theorems applied to asymptotic enumeration II: Multivariate generating functions*, J. Combin. Theory Ser. A **34** (1983), 255–265.

[3] Andrew C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. Amer. Math. Soc. **49** (1941), 122–136.

[4] Michael Drmota, *Systems of functional equations*, Random Structures Algorithms **10** (1997), no. 1–2, 103–124.

[5] ______, *Random trees*, SpringerWienNewYork, 2009.

[6] Christopher Eagle, Zhicheng Gao, Mohamed Omar, Daniel Panario, and Bruce Richmond, *Distribution of the number of encryptions in revocation schemes for stateless receivers*, Fifth Colloquium on Mathematics and Computer Science, Discrete Math. Theor. Comput. Sci. Proc., AI, 2008, pp. 195–206.

[7] Carl-Gustav Esseen, *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law*, Acta Math. **77** (1945), 1–125.

[8] Philippe Flajolet and Robert Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.

[9] Nicko G. Gamkrelidze, *A multidimensional generalization of Esseen's inequality for distribution functions*, Teor. Verojatnost. i Primenen. **22** (1977), no. 4, 897–900.

[10] ——, *A multidimensional generalization of Esseen's inequality for distribution functions*, Theory Probab. Appl. **22** (1977), 877–880, English Translation of the paper in Teor. Verojatnost. i Primenen.

[11] Allan Gut, *Probability: A graduate course*, Springer Texts in Statistics, Springer-Verlag New York, 2005.

[12] Clemens Heuberger, *Hwang's quasi-power-theorem in dimension two*, Quaest. Math. **30** (2007), 507–512.

[13] Clemens Heuberger, Sara Kropf, and Stephan Wagner, *Variances and covariances in the central limit theorem for the output of a transducer*, European J. Combin. **49** (2015), 167–187.

[14] Clemens Heuberger and Helmut Prodinger, *Analysis of alternative digit sets for nonadjacent representations*, Monatsh. Math. **147** (2006), 219–248.

[15] ——, *The Hamming weight of the non-adjacent-form under various input statistics*, Period. Math. Hungar. **55** (2007), 81–96.

[16] Hsien-Kuei Hwang, *On convergence rates in the central limit theorems for combinatorial structures*, European J. Combin. **19** (1998), 329–343.

[17] Sara Kropf, *Variance and covariance of several simultaneous outputs of a Markov chain*, arXiv:1508.05754 [math.CO], 2015.

[18] *The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org`, 2015.

[19] Svetlana Poznanović and Christine E. Heitsch, *Asymptotic distribution of motifs in a stochastic context-free grammar model of RNA folding*, J. Math. Biol. **69** (2014), no. 6, 1743–1772.

[20] Bhagavatula Lakshmi Surya Prakasa Rao, *Another Esseen-type inequality for multivariate probability density functions*, Statist. Probab. Lett. **60** (2002), no. 2, 191–199.

[21] George G. Roussas, *An Esseen-type inequality for probability density functions, with an application*, Statist. Probab. Lett. **51** (2001), no. 4, 397–408.

[22] S. M. Sadikova, *On two-dimensional analogs of an inequality of Esseen and their application to the central limit theorem*, Teor. Verojatnost. i Primenen. **11** (1966), 369–380.

[23] ______, *On two-dimensional analogues of an inequality of Esseen and their application to the central limit theorem*, Theory Probab. Appl. **XI** (1966), 325–335, English Translation of the paper in Teor. Verojatnost. i Primenen.

[24] Richard P. Stanley, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012.