Journal of Theoretical Probability, Vol. 18, No. 2, April 2005 (© 2005)
DOI: 10.1007/s10959-005-3509-6

Generalized Symmetric Polynomials and an Approximate De Finetti Representation

Sergey G. Bobkov[1]

Received October 5, 2003; revised July 11, 2004

For probability measures on product spaces which are symmetric under permutations of coordinates, we study the rate of approximation by mixtures of product measures.

KEY WORDS: Symmetric polynomials; quadratic; induction; de Finetti representation

1. INTRODUCTION

A sequence of random variables $X=(X_1,\dots,X_k)$ on a probability space $(\Omega,\mathscr{A},\mathbf{P})$ with values in a measurable space (M, $\mathscr{B}$) is called exchangeable if its distribution P_X, being a probability measure on the Cartesian power $(M^k,\mathscr{B}^k)$, is symmetric under permutations of coordinates.

A similar definition applies in case of an infinite sequence $X=\{X_k\}_{k=1}^{\infty}$. In that case, under mild regularity assumptions (e.g., when $(M,\mathscr{B})$ is isomorphic to a Borel subset of the unit interval equipped with the Borel σ-algebra), the distribution P_X admits de Finetti's representation in the form of a mixture

$$P_X=\int_{\Pi}\mu_t^{\infty}\,d\pi(t) \tag{1.1}$$

of product probability measures $\mu_t^{\infty}=\mu_t\otimes\mu_t\otimes\cdots$ on $(M^{\infty},\mathscr{B}^{\infty})$. Here (Π,π) is some probability space and $\{\mu_t\}_{t\in\Pi}$ is some family of marginals with the property that the functions $t\to\mu_t(B)$ are π-measurable for all $B\in\mathscr{B}$; see Refs. 3, 4 and 6 for historical comments and general results.

[1] School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church St. S.E., Minneapolis, MN 55455. E-mail: bobkov@math.umn.edu

This representation allows one to study various asymptotic properties of the sequence X such as the law of large numbers, the central limit theorem, by applying known results about independent random variables; cf. e.g., Refs. 1, 7, and 8.

In the general case of a finite exchangeable sequence $X=(X_1,\ldots,X_k)$, the finite-dimensional analogue of Eq. (1.1),

$$P_X(B)=\int \mu_t^k(B)d\pi(t), \quad B\in\mathscr{B}^k, \tag{1.2}$$

is no longer valid, and in fact the class of distributions on M^k symmetric under permutations of coordinates is much wider. Nevertheless, the equality (1.2) remains to hold in a certain approximate sense provided that X represents the beginning of a larger exchangeable sequence $(X_1,\ldots,X_n)$. In terms of the total variation distance

$$||P-Q||_{\mathrm{TV}}=\sup_{B\in\mathscr{B}^k}|P(B)-Q(B)|,$$

a number of results in this direction has been obtained by Diaconis and Freedman[3] (and partly in Ref. 5). It was shown in particular that, for some mixture Q_X of product probability measures on M^k,

$$||P_X-Q_X||_{\mathrm{TV}}\leqslant 1-\frac{n!}{n^k(n-k)!}. \tag{1.3}$$

They also showed that this bound cannot be improved. Actually, if an exchangeable extension $X_1,\ldots,X_n$ exists on the same probability space $(\Omega,\mathscr{A},\mathbf{P})$, one can take $Q_X(B)=\int\mu_\omega^k(B)d\mathrm{P}(\omega)$, that is, with

$$\Pi=\Omega,\quad \pi=\mathbf{P},\quad \mu_\omega=\frac{\delta_{X_{1(\omega)}}+\cdots+\delta_{X_{n(\omega)}}}{n},$$

where δ_x denotes a point mass at x. We will refer to this measure Q_X as associated to P_X.

By letting $n\to\infty$ with fixed k, inequality (1.3) implies de Finetti's theorem for infinite exchangeable sequences. However, for finite exchangeable sequences, the right hand side of Eq. (1.3) is getting small only in the range $k=o(\sqrt{n})$, and is becoming of order 1 for larger values of k. This "negative" observation inspires to look for different distances or conditions under which one is able to judge on the closeness of Q_X to P_X in the whole range $k=o(n)$. In particular, it was proved already in the same paper[3] that for finite spaces M the estimate (1.3) may be complemented with

$$||P_X-Q_X||_{\mathrm{TV}}\leqslant \mathrm{card}(M)\frac{k}{n}.$$

To study the general case, we propose to consider the closeness Q_X to P_X in terms of a ("product variation") distance, defined as

$$||P - Q||_{\mathrm{PV}} = \sup |P(B) - Q(B)|,$$

where P, Q are arbitrary probability measures on M^k, and where the supremum is running over all Cartesian products $B = B_1 \times \cdots \times B_k$ with $B_i \in \mathscr{B}$. In the space of all probability measures on M^k, this distance generates a topology with convergence which is, of course, weaker than the one in the total variation norm. However, when M itself is a metric space inducing a product topology in M^k, the convergence in the product variation norm is stronger than the usual weak convergence of probability distributions. Thus, specializing, for example, to the most interesting space $M = \mathbf{R}$, potentially one will be able to control some canonical distances between probability distribution on $\mathbf{R}^k$, like the Levý or Prokhorov distances, which are responsible for the weak convergence.

Let $\mathscr{F}_k$ denote the collection of all complex-valued functions of f on M^k representable as

$$f(x) = f_1(x_1), \ldots, f_k(x_k), \quad x = (x_1, \ldots, x_k) \in M^k,$$

for some measurable f_i on M such that $|f_i| \leqslant 1$, $1 \leqslant i \leqslant k$.

Theorem 1.1. Assume a sequence $X = (X_1, \ldots, X_k)$ of random variables in $(M, \mathscr{B})$ has an exchangeable extension of length $n \geqslant k$. Then, for any $f \in \mathscr{F}_k$,

$$\left| \int f \, dP_X - \int f \, dQ_X \right| \leqslant C \frac{k}{n}, \tag{1.4}$$

where Q_X is the associated distribution on M^k, and C is a universal constant. In particular,

$$||P_X - Q_X||_{\mathrm{PV}} \leqslant C \frac{k}{n}. \tag{1.5}$$

Inequality (1.5) explicitly measures closeness of $P_X(B)$ to $Q_X(B)$ only for product sets B in M^k. The point of generalization is that, being stated for complex-valued functions, inequality (1.4) contains implicitly information on the closeness of $P_X(B)$ to $Q_X(B)$ for half-spaces $B = \{x \in \mathbf{R}^k : a_1 x_1 + \cdots + a_k x_k \leqslant c\}$. Indeed, we may apply Eq. (1.4) to the family of functions

$$f(x) = e^{it(a_1 x_1 + \cdots + a_k x_k)}, \quad t \in \mathbf{R},$$

and therefore compare the characteristic functions of the weighted sums with respect to P_X and Q_X, respectively. Furthermore, under the product measures, satisfying mild moment conditions, the function $x \to S_k = \frac{x_1+\cdots+x_k}{\sqrt{k}}$ is asymptotically normal, so the inequality (1.4) can be used to study the asymptotic normality of S_k under P_X; for further discussions, cf. (Ref. 2) where a particular case of Eq. (1.4) is treated. Interesting examples come from Convex Geometry when the random vector X is uniformly distributed over a high-dimensional symmetric convex body or represents its high-dimensional projections. Typically, in such examples there is no infinite exchangeable extension. We do not consider this line of applications, but mainly focus on the proof of Theorem 1.1.

The paper is organized as follows. In Section 2, we formulate Theorem 1.1 in terms of the so-called elementary symmetric polynomials. An induction argument is described in Section 3. It reduces the proof of Theorem 1.1 to a certain analytic inequality stated in Lemma 3.1. Sections 4 and 5 are devoted to the proof of this lemma, only.

2. GENERALIZED SYMMETRIC POLYNOMIALS

If $X=(X_1,\dots,X_k)$ has an exchangeable extension $\bar{X}=(X_1,\dots,X_n)$ in M^n, the distribution P_X of the first k coordinates satisfies

$$\int f\,dP_X = \int_{M^n}\left(\int_{M^k} f\,d\mu_x\right)dP_{\bar{X}}(x),$$

where $P_{\bar{X}}$ denotes the distribution of the extended sequence, and where, for a point $x=(x_1,\dots,x_n)\in M^n$, we put

$$\mu_x = \frac{(n-k)!}{n!}\left(\delta_{x_{i_1}}+\cdots+\delta_{x_{i_k}}\right),$$

with summation over all distinct indices $i_1,\dots,i_k$ varying from 1 to n. Similarly, in terms of the product measures $\nu_x = \left(\frac{\delta_{x_1}+\cdots+\delta_{x_n}}{n}\right)^{\otimes k}$ on M^k,

$$\int f\,dQ_X = \int_{M^n}\left(\int_{M^k} f\,d\nu_x\right)dP_{\bar{X}}(x).$$

Therefore,

$$\left|\int f\,dP_X - \int f\,dQ_X\right| \leqslant \int_{M^n}\left|\int_{M^k} f\,d\mu_x - \int_{M^k} f\,d\nu_x\right| dP_{\bar{X}}(x),$$

so inequality (1.4) reduces to the simple "urn model" where one needs to show that

$$\left|\int f\,d\mu_x - \int f\,d\nu_x\right| \leqslant C\frac{k}{n},$$

for any $f \in \mathscr{F}_k$ and all points $x \in M^n$. Moreover, for $f(y)=f_1(y_1), \dots, f_k(y_k)$, the above inequality takes the form

$$\left| \frac{(n-k)!}{n!} \sum f_1(x_{i_1}), \dots, f_k(x_{i_k}) - \prod_{i=1}^{k} \frac{f_i(x_1) + \cdots + f_i(x_n)}{n} \right| \leqslant C \frac{k}{n} \tag{2.1}$$

with summation over all distinct indices $i_1, \dots, i_k$ from 1 to n.

Note that inequality (1.5) amounts to Eq. (2.1) with indicators functions $f_i = 1_{B_i}$ in which case we obtain a certain combinational problem on estimating (both from above and below) the cardinality of a product set with removed points containing repeated coordinates. This seems to be easier task, but actually it leads to the same difficulties we meet in the general case of arbitrary functions f_i.

In the general case, we are dealing with an arbitrary $k \times n$ matrix z with complex-valued entries $z_{ij} = f_i(x_j)$ such that $|z_{ij}| \leqslant 1$. Put

$$\sigma_{n,k}(z) = \frac{(n-k)!}{n!} \sum z_{1i_1} \cdots z_{ki_k}, \tag{2.2}$$

$$s_{n,k}(z) = \prod_{i=1}^{k} \frac{z_{i1} + \cdots + z_{in}}{n}, \tag{2.3}$$

with restrictions for indices $i_1, \dots, i_k$ as before. When the elements in each row of z are equal to each other, $\sigma_{n,k}(z)$ turn into the usual elementary symmetric polynomials of degree k in n complex variables (up to a normalizing factor), cf. (Ref. 9). Therefore, $\sigma_{n,k}$, introduced in Eq. (2.2), may be viewed as generalized elementary symmetric polynomials.

Thus, inequality (1.4) of Theorem 1.1 is equivalent to:

Theorem 2.1. For any $k \times n$ matrix $z = (z_{ij})$ with $|z_{ij}| \leqslant 1$,

$$|\sigma_{n,k}(z) - s_{n,k}(z)| \leqslant C \frac{k}{n},$$

where C is a universal constant.

Note that the left hand side in Eq. (2.3) is vanishing for $k = 1$, so we may and do always assume that $n \geqslant k \geqslant 2$.

3. INDUCTION ARGUMENT

From now on, let us also make the conversion that indices i, t, u, v are reserved to vary from 1 to k, while j and s will vary from 1 to n.

Given a $k\times n$ matrix z, let z^{ij} denote the $(k-1)\times(n-1)$ matrix with entries

$$(z^{ij})_{ts}=z_{ts},\quad t\neq i,\ \ s\neq j.$$

That is, z^{ij} is obtained from z by removing ith row and jth column. Then, from definitions (2.2) and (2.3),

$$\sigma_{n,k}(z)=\frac{1}{n}\sum_{j=1}^{n} z_{ij}\sigma_{n-1,k-1}(z^{ij}),$$
$$s_{n,k}(z)=\frac{1}{n}\sum_{j=1}^{n} z_{ij}\prod_{t\neq i}\frac{z_{t1}+\cdots+z_{tn}}{n},$$

so

$$\sigma_{n,k}(z)-s_{n,k}(z)=\frac{1}{n}\sum_{j=1}^{n} z_{ij}(\sigma_{n-1,k-1}(z^{ij})-s_{n-1,k-1}(z^{ij}))$$
$$+\frac{1}{n}\sum_{j=1}^{n} z_{ij}\left(s_{n-1,k-1}(z^{ij})-\prod_{t\neq i}\frac{z_{t1}+\cdots+z_{tn}}{n}\right).$$

This representation holds true for any fixed value $i\leqslant k$. Averaging over all i with weight $1/k$, we arrive at

$$\sigma_{n,k}(z)-s_{n,k}(z)=\frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n} z_{ij}\left(\sigma_{n-1,k-1}(z^{ij})-s_{n-1,k-1}(z^{ij})\right)$$
$$+\frac{\Sigma_{n,k}(z)}{kn},\tag{3.1}$$

where

$$\Sigma_{n,k}(z)=\sum_{i=1}^{k}\sum_{j=1}^{n} z_{ij}\left(s_{n-1,k-1}(z^{ij})-\prod_{t\neq i}\frac{z_{t1}+\cdots+z_{tn}}{n}\right).\tag{3.2}$$

Let $A_{n,k}$ denote the supremum of $|\sigma_{n,k}(z)-s_{n,k}(z)|$ over all admissible choices of z. Then, in accordance with Eq. (3.1) and using $|z_{ij}|\leqslant 1$,

$$|\sigma_{n,k}(z)-s_{n,k}(z)|\leqslant A_{n-1,k-1}+\frac{1}{kn}|\Sigma_{n,k}(z)|\tag{3.3}$$

Thus, there is a good reason to try to prove Theorem 2.1 by induction, by stating a suitable bound on $|\Sigma_{n,k}(z)|$. What we need is:

Lemma 3.1. In the range $2 \leqslant k \leqslant \frac{n}{8}$, for any $k \times n$ matrix $z = (z_{ij})$ with $|z_{ij}| \leqslant 1$, we have

$$|\Sigma_{n,k}(z)| \leqslant 5k.$$

Proof of Theorem 2.1. Note that $A_{n,k} \leqslant 2$, so $A_{n,k} \leqslant 16\frac{k}{n}$ in the range $k \geqslant \frac{n}{8}$. In the other case, by Lemma 3.1 and using Eq. (3.3), we get

$$A_{n,k} \leqslant A_{n-1,k-1} + \frac{5}{n}.$$

Repeating this inequality $k-1$ times, we arrive at

$$A_{n,k} \leqslant \frac{5}{n} + \frac{5}{n-1} + \cdots + \frac{5}{n-k+2} + A_{n-k+1,1} \leqslant \frac{5(k-1)}{n-k+2} \leqslant 5,$$

since $A_{m,1} = 0$. This completes the proof of Theorems 1.1–2.1 with $C = 16$.

4. QUADRATIC APPROXIMATION. LINEAR TERM

To proof Lemma 3.1, we need some preparations. First let us simplify the expression in Eq. (3.2) by introducing

$$a_i = \frac{z_{i1} + \cdots + z_{in}}{n} \quad \text{and} \quad \varepsilon_{ij} = \frac{a_i - z_{ij}}{n-1},$$

so that

$$s_{n-1,k-1}(z^{ij}) \equiv \prod_{t \neq i} \frac{(z_{t1} + \cdots + z_{tn}) - z_{tj}}{n-1} = \prod_{t \neq i} \frac{na_t - z_{tj}}{n-1} = \prod_{t \neq i} (a_t + \varepsilon_{tj}).$$

Hence, Eq. (3.2) takes a more compact form

$$\Sigma_{n,k}(z) = \sum_{i=1}^{k} \sum_{j=1}^{n} z_{ij} \left(\prod_{t \neq i} (a_t + \varepsilon_{tj}) - \prod_{t \neq i} a_t \right). \tag{4.1}$$

Writing $n\varepsilon_{tj} = \frac{1}{n-1} \sum_{s \neq j} z_{ts} - z_{tj}$ and nothing that the sum over s contains $n-1$ terms, we also obtain that in case $|z_{ij}| \leqslant 1$ (as in Lemma 3.1)

$$|\varepsilon_{tj}| \leqslant \frac{2}{n}. \tag{4.2}$$

In order to estimate the expression (4.1) under (4.2), we apply the following elementary representation holding true for any finite collection of complex numbers (w_r) and (ε_r) such that $|\varepsilon_r| \leqslant \varepsilon$ $(\varepsilon > 0)$:

$$\prod_r (w_r + \varepsilon_r) - \prod_r w_r = \sum_r \left(\varepsilon_r \prod_{u \neq r} w_u \right) + \frac{\theta}{2} \sum_{u \neq v} \left(|\varepsilon_u| |\varepsilon_v| \prod_{r \neq u,v} (|w_r| + \varepsilon) \right).$$

The equality is obtained from Taylor's expansion up to the quadratic term for the function of a real variable, $\varphi(t) = \prod_r (w_r + t\varepsilon_r)$, about the point $t = 0$. Here and throughout, θ denotes a complex number such that $|\theta| \leqslant 1$.

Applying the above equality in Eq. (4.1) with $\varepsilon = \frac{2}{n}$, we obtain the representation

$$\Sigma_{n,k}(z) = L_{n,k}(z) + \frac{\theta}{2} Q_{n,k}(z) \tag{4.3}$$

with linear and quadratic terms with respect to ε_{ij}'s, namely, with

$$L_{n,k}(z) = \sum_{i=1}^{k} \sum_{j=1}^{n} \sum_{t \neq i} z_{ij} \varepsilon_{tj} \prod_{u \neq t,i} a_u,$$

$$Q_{n,k}(z) = \sum_{i=1}^{k} \sum_{j=1}^{n} \sum_{u \neq v; u,v \neq i} |\varepsilon_{uj}| |\varepsilon_{vj}| \prod_{t \neq u,v,i} (|a_t| + \varepsilon).$$

Thus, the proof of Lemma 3.1 is reduced to showing that both $|L_{n,k}(z)|$ and $Q_{n,k}(z)$ are at most Ak. Here we concentrate on the linear term, and postpone the problem regarding the quadratic term to the next section.

For simplicity of notations and without loss in generality, assume $a_i \neq 0$, for all $i \leqslant k$. According to the definition of ε_{tj}, we obtain

$$L_{n,k}(z) = -\frac{a_1, \ldots, a_k}{n-1} \sum_{(t,i): t \neq i} \frac{1}{a_i a_t} \sum_{j=1}^{n} z_{ij} (z_{tj} - a_t),$$

where the first sum is double and is taken over all pairs of distinct indices (t, i) varying from 1 to k. Writing $z_{ij} = (z_{ij} - a_i) + a_i$, we split the sums into

$$\begin{aligned} L_{n,k}(z) &= -\frac{a_1, \ldots, a_k}{n-1} \sum_{t \neq i} \frac{1}{a_i a_t} \sum_{j=1}^{n} (z_{ij} - a_i)(z_{tj} - a_t) \\ &\quad - \frac{a_1, \ldots, a_k}{n-1} \sum_{t \neq i} \frac{1}{a_t} \sum_{j=1}^{n} (z_{ij} - a_t) \end{aligned}$$

and observe that the last sum on the right (over j) is vanishing. Hence,

$$L_{n,k}(z) = -\frac{a_1, \ldots, a_k}{n-1} \sum_{t \neq i} \frac{1}{a_i a_t} \sum_{j=1}^{n} (z_{ij} - a_i)(z_{tj} - a_t).$$

Now, to estimate the inner sum, by Cauchy's inequality,

$$\frac{1}{n} \left| \sum_{j=1}^{n} (z_{ij} - a_i)(z_{tj} - a_t) \right| \leqslant \sqrt{\frac{1}{n} \sum_{j=1}^{n} |z_{ij} - a_i|^2} \sqrt{\frac{1}{n} \sum_{j=1}^{n} |z_{tj} - a_t|^2}. \tag{4.4}$$

But

$$\frac{1}{n} \sum_{j=1}^{n} |z_{ij} - a_i|^2 = \frac{1}{n} \sum_{j=1}^{n} (|z_{ij}|^2 - |a_i|^2) \leqslant 1 - |a_i|^2, \tag{4.5}$$

where we have used the assumption $|z_{ij}| \leqslant 1$ on the last step. Hence,

$$|L_{n,k}(z)| \leqslant \frac{n}{n-1} |a_1|, \ldots, |a_k| \sum_{t \neq i} \frac{1}{|a_i||a_t|} \sqrt{1 - |a_i|^2} \sqrt{1 - |a_t|^2}.$$

What is remarkable is that this bound does not in essense involve n. Moreover, we can think of $|a_i|$ as of arbitrary numbers in (0, 1]. Making the substitution $x_i = \sqrt{\frac{1}{|a_i|^2} - 1}$, the bound takes the form

$$|L_{n,k}(z)| \leqslant \frac{n}{n-1} \frac{\sum_{t \neq i} x_i x_t}{\prod_{i=1}^{k} \sqrt{1 + x_i^2}}. \tag{4.6}$$

Let B_k be the optimal constant in

$$\sum_{t \neq i} x_i x_t \leqslant B_k \prod_{i=1}^{k} \sqrt{1 + x_i^2}, \quad x_1, \ldots, x_k \geqslant 0.$$

For example, when $k=2$, the inequality becomes $2x_1x_2 \leqslant B_2\sqrt{1+x_1^2}\sqrt{1+x_2^2}$, so $B_2=2$. If $k\geqslant 3$,

$$\sum_{t\neq i} x_i x_t = \sum_{t\neq i;t,i<k} x_i x_t + 2x_k \sum_{i=1}^{k-1} x_i \leqslant B_{k-1}\prod_{i=1}^{k-1}\sqrt{1+x_i^2} + 2x_k \sum_{i=1}^{k-1} x_i.$$

By Cauchy's inequality $(\sum_{i=1}^{k-1} x_i)^2 \leqslant (k-1)\sum_{i=1}^{k-1} x_i^2 \leqslant (k-1)\prod_{i=1}^{k-1}(1+x_i^2)$, so

$$\sum_{t\neq i} x_i x_t \leqslant \prod_{i=1}^{k-1}\sqrt{1+x_i^2}(B_{k-1}+2(k-1)x_k).$$

Dividing both sides by $\prod_{i=1}^{k}\sqrt{1+x_i^2}$ and maximizing over x_k, we get a recursive inequality $B_k^2 \leqslant B_{k-1}^2 + 4(k-1)^2$. Applying it successively $k-2$ times, we obtain

$$B_k^2 \leqslant 4(k-1)^2 + 4(k-2)^2 + \cdots + 4\cdot 2^2 + B_2^2 = 2k(k+1) \leqslant 4k^2.$$

Finally, recalling Eq. (4.6) and using $\frac{n}{n-1}\leqslant 2$, we may conclude:

Lemma 4.1. $|L_{n,k}(z)| \leqslant 4k.$

5. QUADRATIC TERM. PROOF OF LEMMA 3.1

Now let's turn to the quadratic term and write it as

$$Q_{n,k}(z) = \frac{1}{(n-1)^2}\sum_{i=1}^{k}\sum_{j=1}^{n}\prod_{t\neq i}(|a_t|+\varepsilon)\sum_{u\neq v;u,v\neq i}\frac{|z_{uj}-a_u||z_{vj}-a_v|}{(|a_u|+\varepsilon)(|a_v|+\varepsilon)},$$

where $\varepsilon = \frac{2}{n}$. The sum $\sum_{j=1}^{n}|z_{uj}-a_u||z_{vj}-a_v|$ can be estimated as in Eqs. (4.4) and (4.5), so

$$Q_{n,k}(z) \leqslant \frac{n}{(n-1)^2}\sum_{i=1}^{k}\prod_{t\neq i}(|a_t|+\varepsilon)\sum_{u\neq v;u,v\neq i}\frac{\sqrt{1-|a_u|^2}\sqrt{1-|a_v|^2}}{(|a_u|+\varepsilon)(|a_v|+\varepsilon)}.$$

Removing constraints $t\neq i$ and $u,v\neq i$ and performing summation over $i\leqslant k$, we obtain a simpler bound

$$Q_{n,k}(z) \leqslant \frac{2nk}{(n-1)^2}(|a_1|+\varepsilon)\cdots(|a_k|+\varepsilon)\sum_{u<v}\frac{\sqrt{1-|a_u|^2}}{|a_u|+\varepsilon}\frac{\sqrt{1-|a_v|^2}}{|a_v|+\varepsilon},$$

that is,

$$Q_{n,k}(z) \leqslant \frac{2nk}{(n-1)^2} \Psi_k(|a_1|, \ldots, |a_k|), \tag{5.1}$$

where

$$\Psi_k(b_1, \ldots, b_k) = (b_1+\varepsilon) \cdots (b_k+\varepsilon) \sum_{u<v} \frac{\sqrt{1-b_u^2}}{b_u+\varepsilon} \frac{\sqrt{1-b_v^2}}{b_v+\varepsilon}.$$

As a result, we arrived at a quantity where b_i may be thought of as arbitrary numbers in [0, 1]. For such variables, introduce also

$$\psi_k(b_1, \ldots, b_k) = \sum_{u=1}^{k} (b_1+\varepsilon) \cdots (b_{u-1}+\varepsilon) \sqrt{1-b_u^2} (b_{u+1}+\varepsilon) \cdots (b_k+\varepsilon).$$

It should be clear that the two functions are related by

$$\Psi_k(b_1, \ldots, b_k) = \psi_{k-1}(b_1, \ldots, b_{k-1}) \sqrt{1-b_k^2} + \Psi_{k-1}(b_1, \ldots, b_{k-1})(b_k+\varepsilon).$$

Maximizing the right hand side over all $b_k \in [0, 1]$, we obtain a recursive inequality

$$\Psi_k \leqslant \sqrt{\psi_{k-1}^2 + \Psi_{k-1}^2} + \varepsilon \Psi_{k-1}. \tag{5.2}$$

Thus, to get a proper bound $\Psi_k = O(k)$, let us concentrate on estimating the function ψ_k. In view of Eq. (5.2), we need a bound of the from $\psi_k = O(\sqrt{k})$.

Lemma 5.1. If $\frac{k-1}{n} \leqslant \frac{1}{8}$, then $\psi_k \leqslant \sqrt{5.66k}$.

Proof. From definition it follows that

$$\psi_k(b_1, \ldots, b_k) = \sqrt{1-b_k^2} \prod_{i=1}^{k-1} (b_i+\varepsilon) + (b_k+\varepsilon) \psi_{k-1}(b_1, \ldots, b_{k-1}).$$

Since $0 \leqslant b_i \leqslant 1$,

$$\psi_k(b_1, \ldots, b_k) \leqslant (1+\varepsilon)^{k-1} \sqrt{1-b_k^2} + (b_k+\varepsilon) \psi_{k-1}(b_1, \ldots, b_{k-1}).$$

Maximizing the right hand side over all b_k, we obtain

$$\psi_k \leqslant \sqrt{(1+\varepsilon)^{2(k-1)} + \psi_{k-1}^2} + \varepsilon \psi_{k-1}. \tag{5.3}$$

Now, it is easy to perform induction on k in order to prove that

$$\psi_k \leqslant \sqrt{C^2 k}. \tag{5.4}$$

We have $\psi_1(b_1) = \sqrt{1-b_1^2} \leqslant 1$, so Eq. (5.4) holds with $C=1$ for $k=1$. For $k \geqslant 2$, using $(1+\varepsilon)^{2(k-1)} \leqslant (1+\frac{2}{n})^{n/4} \leqslant \sqrt{e}$, we derive from Eq. (5.3) and the induction hypothesis that

$$\psi_k \leqslant \sqrt{\sqrt{e} + C^2(k-1)} + C\varepsilon\sqrt{k-1}.$$

Dividing by C and squaring, we see that the right hand side does not exceed $C\sqrt{k}$ if and only if

$$\frac{\sqrt{e}}{C^2} + \varepsilon^2(k-1) + 2\varepsilon\sqrt{k-1}\sqrt{\frac{\sqrt{e}}{C^2} + (k-1)} \leqslant 1.$$

The last square root dominates the pre-last one, so the above inequality would follow from

$$\frac{\sqrt{e}}{C^2} + \varepsilon^2(k-1) + 2\varepsilon\left(\frac{\sqrt{e}}{C^2} + (k-1)\right) \leqslant 1. \tag{5.5}$$

Now, since $n \geqslant 8(k-1) \geqslant 8$, we have $\varepsilon^2(k-1) \leqslant \left(\frac{2}{n}\right)^2 \frac{n}{8} \leqslant \frac{1}{16}$ and $2\varepsilon(k-1) \leqslant \frac{1}{2}$. Hence, Eq. (5.5) is implied by $\frac{3}{2}\frac{\sqrt{e}}{C^2} \leqslant \frac{7}{16}$, that is, by $C^2 \geqslant \frac{24}{7}\sqrt{e} \approx 5.65$.

Lemma 5.2. If $\frac{k}{n} \leqslant \frac{1}{8}, k \geqslant 2$, then $\Psi_k \leqslant 2k$.

Proof. Now, we perform induction on k in order to prove the inequality

$$\Psi_k \leqslant Bk$$

in the indicated range of k. Recall that $\Psi_2(b_1, b_2) = \sqrt{1-b_1^2}\sqrt{1-b_2^2} \leqslant 1$, so $B=1$ works well for $k=2$. If $k \geqslant 3$, we derive from Eq. (5.2), Lemma 5.2, and the induction hypothesis that

$$\Psi_k \leqslant \sqrt{C^2(k-1) + B^2(k-1)^2} + B\varepsilon(k-1).$$

Hence our task is to show that

$$\sqrt{C^2(k-1) + B^2(k-1)^2} + B\varepsilon(k-1) \leqslant Bk.$$

Dividing by B and putting $\gamma = \frac{C^2}{B^2}$, the inequality turns into

$$\sqrt{\gamma(k-1)+(k-1)^2} \leqslant (1-\varepsilon)k+\varepsilon.$$

Squaring and rearranging, we are reduced to $\varepsilon(2+\varepsilon)k^2 \leqslant (2+2\varepsilon-2\varepsilon^2-\gamma)k+\gamma+\varepsilon^2$. Omitting the terms $2\varepsilon-2\varepsilon^2, \gamma+\varepsilon^2$, dividing by k, and using $\varepsilon k \leqslant \frac{1}{4}$, we obtain stronger inequality

$$\frac{1}{4}(2+\varepsilon) \leqslant 2-\gamma.$$

Due to the assumption $n \geqslant 8k$ and since $k \geqslant 3$, we get $\varepsilon = \frac{2}{n} \leqslant \frac{1}{12}$, so one can take $\gamma = 2-\frac{1}{4}(2+\frac{1}{12}) = \frac{3}{2}-\frac{1}{48}$. Thus, we may take $B^2 = \frac{C^2}{\gamma} < 3.85$.

Proof of Lemma 3.1. Recalling the bound (5.1) on the quadratic term $Q_{n,k}$, Lemma 5.2 implies that

$$Q_{n,k}(z) \leqslant \frac{4nk^2}{(n-1)^2} \leqslant k,$$

whenever $\frac{k}{n} \leqslant \frac{1}{8}, k \geqslant 2$. It remains to combine this bound with Lemma 4.1. Then, by the representation (3.2),

$$|\Sigma_{n,k}(z)| \leqslant |L_{n,k}(Z)| + \frac{1}{2}Q_{n,k}(z) \leqslant 5k.$$

ACKNOWLEDGMENT

Research was partially supported by NSF Grant DMS-0405587.

REFERENCES

1. Blum, J. R., Chernoff, H., Rosenblatt, M., and Teicher, H. (1958). Central limit theorems for interchangeable processes. *Can. J. Math.* **10**, 222–229.
2. Bobkov, S. G. (2004). Concentration of normalized sums and a central limit theorem for non-correlated random variables. *Ann. Probab.* **32**(4), 2884–2907.
3. Diaconis, P., and Freedman, D. (1980). Finite exchangeble sequences. *Ann. Probab.* **8**(4), 745–764.
4. Dubins, L. E., and Freedman, D. (1979). Exchangeable processes need not be mixtures of independent, identically distributed random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **48**(2), 115–132.
5. Freedman, D. (1977). A remark on the difference between sampling with and without replacement. *J. Amer. Stat. Assoc.* **73**, 681.
6. Hewitt, E., and Savage, L. J. (1955). Symmetric measures on Cartesian spaces. *Trans. Amer. Math. Soc.* **80**, 470–501.
7. Jiang, X., and Hahn, M.G. (2002). Emprical central limit theorems for exchangeable random variables. *Stat. Probab. Lett.* **59**, 75–81.

8. Klass, M., and Teicher, H. (1987). The central limit theorem for exchangeable random variables without moment assumptions. *Ann. Probab.* **15**(1), 138–153.
9. Prasolov, V. V. (2004). *Polynomials*, Translated from the 2001 Russian 2nd edn., by Dimitry Leites, Algorithms and Computation in Mathematics, Vol. 11, Springer-Verlag, Berlin.