JOURNAL OF MULTIVARIATE ANALYSIS 11, 581–598 (1981)

Representations for Partially Exchangeable Arrays of Random Variables

DAVID J. ALDOUS

University of California, Berkeley

Communicated by D. A. Fraser

Consider an array of random variables $(X_{i,j})$, $1 \leqslant i,j < \infty$, such that permutations of rows or of columns do not alter the distribution of the array. We show that such an array may be represented as functions $f(\alpha, \xi_i, \eta_j, \lambda_{i,j})$ of underlying i.i.d. random variables. This result may be useful in characterizing arrays with additional structure. For example, we characterize random matrices whose distribution is invariant under orthogonal rotation, confirming a conjecture of Dawid.

1. INTRODUCTION

A sequence $\mathbf{Z} = (Z_i)_{i \geqslant 1}$ is called *exchangeable* if for each $n \geqslant 1$ and each permutation π of $\mathbb{N}$,

$$(Z_1, ..., Z_n) \sim (Z_{\pi(1)}, ..., Z_{\pi(n)}), \tag{1.1}$$

where $\sim$ denotes equality in distribution. A classical theorem, essentially due to de Finetti, characterizes infinite exchangeable sequences as mixtures of i.i.d. sequences. In this paper we investigate related ideas for arrays of random variables.

Let $\mathbf{X} = (X_{i,j})_{1 \leqslant i,j < \infty}$ be an array. Let R_i denote the ith row, that is, $R_i = (X_{i,j}: j \geqslant 1)$. Call $\mathbf{X}$ *row-exchangeable* if the sequence $(R_i)_{i \geqslant 1}$ of random vectors is exchangeable. Let $C_j = (X_{i,j}: i \geqslant 1)$ denote the jth column. Call $\mathbf{X}$ *column-exchangeable* if $(C_j)_{j \geqslant 1}$ is exchangeable. Finally, call $\mathbf{X}$ *row and column exchangeable* (RCE) if it is both row-exchangeable and column-exchangeable.

Our main result is the characterization of RCE arrays given in Theorem

Received May 11, 1979; revised March 9, 1981.

AMS 1970 subject classification: Primary 60G99; Secondary 62E10.

Key words and phrases: Exchangeability, spherical matrices.

0047-259X/81/040581-18$02.00/0

1.4 below. Throughout the paper, let α, $(\xi_i)_{i \geq 1}$, $(\eta_j)_{j \geq 1}$ and $(\lambda_{i,j})_{i,j \geq 1}$ be random variables which are

(i) mutually independent,
(ii) distributed uniformly on $[0, 1]$. (1.2)

Given such random variables, let

$$X^*_{i,j} = f(\alpha, \xi_i, \eta_j, \lambda_{i,j}) \tag{1.3}$$

for some measurable function f. Clearly $\mathbf{X}^*$ is a RCE array. Call an array of this form a *standard* RCE array.

THEOREM 1.4. *Let* $\mathbf{X}$ *be a RCE array. Then there exists a standard RCE array* $\mathbf{X}^*$ *such that* $\mathbf{X}^* \sim \mathbf{X}$.

This result seems of some interest in its own right. Moreover, it gives a new perspective on de Finetti's theorem, which we now discuss briefly (for modern treatments see Kingman [12], Dynkin [7], and Dellacherie and Meyer [4]). The informal concept of an exchangeable sequence (Z_i) being a "mixture of i.i.d. sequences" is usually formalized in one of two ways. First, by conditioning: let $\mathscr{T}$ denote the tail σ-field of (Z_i), and then conditional on $\mathscr{T}$ the variables $Z_1, Z_2, \ldots$ are i.i.d. Or second, by Choquet theory: identifying random sequences with their distributions, the class of exchangeable sequences is a convex set whose extreme points are the i.i.d. sequences, and each exchangeable sequence has an integral representation in terms of these extreme points. Now with these usual forms, the analogy between de Finetti's theorem and Theorem 1.4 is opaque. To make it transparent, we need an unusual form of the Finetti's theorem.

LEMMA 1.5. *An infinite sequence* $\mathbf{Z} = (Z_i)_{i \geq 1}$ *is exchangeable if and only if there exists a measurable function* f *such that* $\mathbf{Z} \sim \mathbf{Z}^* = (f(\alpha, \xi_i))_{i \geq 1}$.

The paper is organized as follows. Section 2 contains the proofs of Lemma 1.5 and Theorem 1.4, preceded by the technical machinery necessary. Section 3 contains further results relating the structure of RCE arrays to their functional representations. In Section 4 these results are applied to solve a problem posed by Dawid [2]. Given an array $\mathbf{Y}$, consider for each n the subarray $Y^n = (Y_{i,j})_{1 \leq i,j \leq n}$ as a $(n \times n)$ matrix. Dawid [1] calls $\mathbf{Y}$ *spherical* if for each n

$$U_1 Y^n U_2 \sim Y^n \qquad \text{for all orthogonal } (n \times n) \text{ matrices } U_1, U_2. \tag{1.6}$$

Theorem 4.3 shows that spherical arrays with square-integrable entries have representations involving matrices with Normal entries. Finally, Section 5 describes recent related results.

2. Proof of the Main Result

Everything through Corollary 2.5 is standard, though unfortunately there is no standard terminology. Let us first review some measure theory. Two measurable spaces are *Borel-isomorphic* if there exists a bijection ϕ between them such that ϕ and ϕ^{-1} are measurable. A *Borel space* is a measurable space which is Borel-isomorphic to some Borel subset of the real line.

As in the Introduction, α, ξ_i, η_j, $\lambda_{i,j}$ denote independent r.v.'s distributed uniformly on $[0, 1]$. Constructing families of r.v.'s with prescribed distributions from such r.v.'s is called *coding*.

Lemma 2.1 (Coding Lemma). (a) *Let Y be a r.v. taking values in a Borel space S. Then there exists a measurable function f: $[0, 1] \to S$ such that $Y \sim f(\xi_1)$.*

(b) *Suppose further that $U_1 \sim U_2$ are r.v.'s taking values in a Borel spaces S', and suppose that U_2 is independent of ξ_1. Then there exists a measurable function g: $S' \times [0, 1] \to S$ such that $(U_1, Y) \sim (U_2, g(U_2, \xi_1))$.*

Proof. (a) First suppose $S = \mathbb{R}$. Let $F(y)$ be the distribution function of Y, and let $F^{-1}(x)$ be the inverse function. Then $Y \sim F^{-1}(\xi_1)$. The extention to the Borel space case is an immediate consequence of the definition of Borel space.

(b) Again, first suppose $S = \mathbb{R}$. Let $F(y|u)$ be the conditional distribution function of Y given U_1. For each u let $F^{-1}(x|u)$ be the inverse function. Then $(U_1, Y) \sim (U_2, F^{-1}(\xi_1 | U_2))$. Again, the extension to the case where S is a Borel space is an immediate consequence of the definition of Borel space.

Recall that sequence space $\mathbb{R}^\infty$ is a Borel space ([16, A.7]). A sequence $(X_i)_{i \geqslant 1}$ of real-valued r.v.'s may be regarding as a single r.v. $\mathbf{X}$ taking values in the Borel space $\mathbb{R}^\infty$. So the Coding Lemma implies that a sequence $\mathbf{X} = (X_i)$ may be coded as $\mathbf{X} \sim f(\xi_1)$ for some f: $[0, 1] \to \mathbb{R}^\infty$. Similarly, all the results in this section stated for individual r.v.'s may be applied to sequences or arrays of r.v.'s: we use this technique without further comment.

We make extensive use of conditional independence. For the formal definitions, let ϕ and ϕ_i denote bounded measurable functions. Call a family $\{X_i: i \in I\}$ *conditionally independent* (c.i.) given $\mathscr{G}$ if $E(\prod \phi_j(X_{i_j}) | \mathscr{G}) = \prod E(\phi_j(X_{i_j}) | \mathscr{G})$ for all finite subsets $\{i_1, ..., i_k\}$ of I and all $\phi_1, ..., \phi_k$. In this definition we identify random variables with the σ-fields they generate, so, for example, $\{\mathscr{F}_i: i \in I\}$ are c.i. given Z if $E(\prod F_{i_j} | Z) = \prod E(F_{i_j} | Z)$ for bounded $\mathscr{F}_{i_j}$-measurable r.v.'s F_{i_j}.

Lemma 2.2 ([3, Theorem 1.45]). *A pair $\{X, Y\}$ are c.i. given $\mathscr{G}$ if and only if $E(\phi(X) | Y, \mathscr{G}) = E(\phi(X) | \mathscr{G})$ for each ϕ.*

Call a family $\{X_i: i \in I\}$ *conditionally identically distributed* (c.i.d.) given $\mathscr{G}$ if $E(\phi(X_i)|\mathscr{G}) = E(\phi(X_j)|\mathscr{G})$ for each ϕ, i, j. When $\mathscr{G} = \sigma(Y)$ this is equivalent to $(X_i, Y) \sim (X_j, Y)$ for all i, j.

Remark. The reader should think of assertions "X and Y are c.i. (or c.i.d.) given $\mathscr{G}$" in intuitive terms as "once you know $\mathscr{G}$, then X and Y are independent (or identically distributed)."

The distribution of a random vector with independent components is obviously determined by the marginal distributions. Here is the "conditional" form of this result. The easy proof is omitted.

LEMMA 2.3 (Identification Lemma). *Let* $\mathbf{Y} = (Y_i: i \in I)$, $\mathbf{Y}^* = (Y_i^*: i \in I)$, *and* $\mathscr{G}$ *be such that*

(i) $\{Y_i: i \in I\}$ *are c.i. given* $\mathscr{G}$,

(ii) $\{Y_i^*: i \in I\}$ *are c.i. given* $\mathscr{G}$,

(iii) *for each* i, $\{Y_i, Y_i^*\}$ *are c.i.d. given* $\mathscr{G}$.

Then $\mathbf{Y} \sim \mathbf{Y}^*$.

We now give a precise statement of the most familiar form of de Finetti's theorem and a related fact (see Kingman [12]).

PROPOSITION 2.4. *Let* $\{Z_1, Z_2, \ldots\}$ *be exchangeable, with tail* σ*-field* $\mathscr{T}$. *Then* $\{Z_i: i \geqslant 1\}$ *are c.i. and c.i.d. given* $\mathscr{T}$. *And* $(1/n)\sum_{i=1}^n \phi(Z_i) \to E(\phi(Z_1)|\mathscr{T})$ *a.s. for any bounded* ϕ.

We shall need the analogue for doubly-infinite sequences.

COROLLARY 2.5. *Let* $\{\ldots, Z_{-1}, Z_0, Z_1, \ldots\}$ *be exchangeable, and let* $\mathscr{T}_- = \bigcap_{n > -\infty} \sigma(Z_i: i \leqslant n)$. *Then*

(a) $\{Z_i: -\infty < i < \infty\}$ *are c.i. and c.i.d. given* $\mathscr{T}_-$.

(b) $E(\phi(Z_1)|\mathscr{T}_-)$ *is essentially* $\bigcap_{n > -\infty} \sigma(\phi(Z_i): i \leqslant n)$*-measurable.*

Proof. Apply Proposition 2.4 to $\{Z_k, Z_{k-1}, \ldots\}$.

This concludes the technical preliminaries. We now start proving the results of the present paper.

Proof of Lemma 1.5. Without loss of generality we may suppose that $\{Z_1, Z_2, \ldots\}$ is part of a doubly-infinite exchangeable sequence $\mathbf{Z} = \{\ldots, Z_{-1}, Z_0, Z_1, \ldots\}$, since we could replace (Z_i) by $(Z_i^*) = (Z_{\theta(i)})$ for some bijection $\theta: \mathbb{Z} \to \mathbb{N}$. Let α^* denote the vector $(\ldots, Z_{-1}, Z_0)$. Then

$$\{Z_i: i \geqslant 1\} \text{ are c.i. given } \alpha^* \tag{2.6}$$

(the proof of (2.6) is deferred). Now take α, ξ_i as in (1.2), independent of $\mathbf{Z}$. By the Coding Lemma there exists g such that $(\alpha^*, g(\alpha^*, \xi_1))) \sim (\alpha^*, Z_1)$. Then $(\alpha^*, g(\alpha^*, \xi_i)) \sim (\alpha^*, Z_i)$ for each $i \geqslant 1$ since plainly neither distribution depends on i. Since $\{g(\alpha^*, \xi_i): i \geqslant 1\}$ are plainly c.i. given α^*, the Identification Lemma and (2.6) imply

$$(g(\alpha^*, \xi_i))_{i \geqslant 1} \sim (Z_i)_{i \geqslant 1}. \tag{2.7}$$

Now code α^* as $h(\alpha)$. Then $(\alpha^*, \xi_1, \xi_2, ...) \sim (h(\alpha), \xi_1, \xi_2, ...)$, and so $(g(\alpha^*, \xi_i))_{i \geqslant 1} \sim (g(h(\alpha), \xi_i))_{i \geqslant 1}$. Thus, Lemma 1.5 holds for $f(\cdot, \cdot) = g(h(\cdot), \cdot)$.

The reader should see, at least intuitively, that (2.6) is a consequence of de Finetti's theorem. Here is the formal argument. Let $\mathscr{E}_-$ be as in Corollary 2.5. Then

$$\begin{aligned} E(\phi_1(Z_1)\,\phi_2(Z_2)\,|\,\alpha^*) &= E(\phi_1(Z_1)\,\phi_2(Z_2)\,|\,\alpha^*, \mathscr{E}_-) \qquad \text{as} \quad \mathscr{E}_- \subset \sigma(\alpha^*) \\ &= E(\phi_1(Z_1)\,\phi_2(Z_2)\,|\,\mathscr{E}_-) \end{aligned} \tag{2.8}$$

by Corollary 2.5(a) and Lemma 2.2;

$$= \prod E(\phi_i(Z_i)\,|\,\mathscr{E}_-) \qquad \text{by Corollary 2.5(a).}$$

Similarly,

$$E(\phi_i(Z_i)\,|\,\alpha^*) = E(\phi_i(Z_i)\,|\,\mathscr{E}_-). \tag{2.9}$$

Thus, $E(\phi_1(Z_1)\,\phi_2(Z_2)\,|\,\alpha^*) = \prod E(\phi_i(Z_i)\,|\,\alpha^*)$, and the same argument works for k-tuples, proving (2.6).

Remarks. The above proof was designed to exhibit the technical machinery at work in a simple setting. The proof of Theorem 1.4 is similar in outline, but (2.6) is replaced by the less obvious relations of Lemma 2.10. Observe that (2.6), (2.8), and (2.9) hold for any exchangeable sequence $\{..., Z_{-1}, Z_0, Z_1, ...\}$. We shall apply these later to sequences of rows, for example.

Proof of Theorem 1.4. As above, we may suppose the given RCE array $(X_{i,j})_{i,j \geqslant 1}$ is part of an extended RCE array $\mathbf{X} = (X_{i,j})_{-\infty < i,j < \infty}$. We shall need the following notation. Let C_j denote the jth column of the extended array, that is, $C_j = (X_{i,j}: i \in \mathbb{Z})$. Put $C_j^+ = (X_{i,j}: i \geqslant 1)$, $C_j^- = (X_{i,j}: i \leqslant 0)$. Let R_i denote the ith row $(X_{i,j}: j \in \mathbb{Z})$, and let $R_i^+ = (X_{i,j}: j \geqslant 1)$, $R_i^- = (X_{i,j}: j \leqslant 0)$. Finally, put

$$\begin{aligned} &A = (X_{i,j}: i,j \leqslant 0), \\ &C_j^* = (A, C_j^-), \qquad R_i^* = (A, R_i^-) \\ &\mathscr{G} = \sigma(X_{i,j}: \min(i,j) \leqslant 0) = \sigma(A, R_i^*, C_j^*: i,j \geqslant 1). \end{aligned}$$

The next lemma gives the conditional independence properties needed for the construction. Its proof is deferred.

LEMMA 2.10. (a) $\{X_{i,j}: i,j \geqslant 1\}$ *are c.i. given* $\mathscr{G}$.

(b) $\{C_1^*, C_2^*, \ldots; R_1^*, R_2^*, \ldots\}$ *are c.i. given* A.

(c) $X_{i,j}$ *and* $\mathscr{G}$ *are c.i. given* $\sigma(A, R_i^*, C_j^*)$, *for each* $i,j \geqslant 1$.

Take α, ξ_i, η_j, $\lambda_{i,j}$ as in (1.2), independent of $\mathbf{X}$. By the Coding Lemma there exists g such that, putting $X_{1,1}^* = g(A, R_1^*, C_1^*, \lambda_{1,1})$, we have $(A, R_1^*, C_1^*, X_{1,1}^*) \sim (A, R_1^*, C_1^*, X_{1,1})$. For each $i,j \geqslant 1$ define

$$X_{i,j}^* = g(A, R_i^*, C_j^*, \lambda_{i,j}). \tag{2.11}$$

Then

$$(A, R_i^*, C_j^*, X_{i,j}^*) \sim (A, R_i^*, C_j^*, X_{i,j}), \qquad \text{each} \quad i,j \geqslant 1 \tag{2.12}$$

because neither distribution depends on (i,j). We shall prove an analogue of (2.7):

$$(X_{i,j}^*)_{i,j \geqslant 1} \sim (X_{i,j})_{i,j \geqslant 1}. \tag{2.13}$$

We wish to get (2.13) from the Identification Lemma. Plainly, $\{X_{i,j}^*: i,j \geqslant 1\}$ are c.i. given $\mathscr{G}$, so in view of Lemma 2.10(a) it is only required to verify that for each (i,j), $\{X_{i,j}, X_{i,j}^*\}$ are c.i.d. given $\mathscr{G}$. But

$$\begin{aligned} E(\phi(X_{i,j})|\mathscr{G}) &= E(\phi(X_{i,j})|A, R_i^*, C_j^*) && \text{by Lemma 2.10(c)} \\ &= E(\phi(X_{i,j}^*)|A, R_i^*, C_j^*) && \text{by (2.12)} \\ &= E(\phi(X_{i,j}^*)|\mathscr{G}) && \text{from (2.11),} \end{aligned}$$

since $\lambda_{i,j}$ is independent of $\mathscr{G}$.

Having established (2.13), the next step is to code R_i^* and C_j^*. By the Coding Lemma there exists h' such that, putting $\xi_1^* = h'(A, \xi_1)$, we have $(A, \xi_1^*) \sim (A, R_1^*)$. For each $i \geqslant 1$ define $\xi_i^* = h'(A, \xi_i)$. Then $(A, \xi_i^*) \sim (A, R_i^*)$ for each i, since the distributions do not depend on i. Similarly, there exists h'' such that $\eta_j^* = h''(A, \eta_j)$ satisfies $(A, \eta_j^*) \sim (A, R_j^*)$ for each $j \geqslant 1$. We now want to use the Identification Lemma to prove

$$(A, \xi_i^*, \eta_j^*)_{i,j \geqslant 1} \sim (A, R_i^*, C_j^*)_{i,j \geqslant 1}. \tag{2.14}$$

In view of Lemma 2.10(b) and the distributional identities above, we need only check that $\{A, \xi_i^*, \eta_j^*: i,j \geqslant 1\}$ are c.i. given A. But this holds because $\{\xi_i, \eta_j: i,j \geqslant 1\}$ are independent of each other and of A.

In the construction above, $(\xi_i^*, \eta_j^*) = h(A, \xi_i, \eta_j)$ for a certain function h. Putting together (2.13), (2.11), and (2.14), we see that the array $(X_{i,j})$ has the same distribution as the array $\hat{X}_{i,j} = g(A, h(A, \xi_j, \eta_j), \lambda_{i,j})$. Finally, code A as a function of α as in the proof of Lemma 1.5, and the desired representation is established.

Proof of Lemma 2.10. The assertions of Lemma 2.10 are consequences of de Finetti's theorem applied to the exchangeable sequences (R_i) and (C_j). Recall the notation $R_i = (R_i^+, R_i^-)$, $C_j = (C_j^+, C_j^-)$. We shall verify conditional independence for pairs ϕ_1, ϕ_2. The arguments for k-tuples are identical. For a σ-field $\sigma(\mathscr{F}_1, \mathscr{F}_2, \mathscr{F}_3)$ call sets of the form $F_1 \times F_2 \times F_3$ $(F_i \in \mathscr{F}_i)$ *generating rectangles.*

(a) We shall show

$$\{R_i^+ : i \geqslant 1\} \qquad \text{are c.i. given } \mathscr{G}. \tag{2.15}$$

Then a symmetric argument shows $\{C_j^+ : j \geqslant 1\}$ are c.i. given $\mathscr{G}$, and (a) follows. To prove (2.15), first apply (2.6) to (R_i), giving

$$\{R_i : i \geqslant 1\} \qquad \text{are c.i. given} \quad \mathscr{U} = \sigma(R_i : i \leqslant 0). \tag{2.16}$$

Next we assert

$$E(\phi_1(R_1^+) | \mathscr{G}) = E(\phi_1(R_1^+) | \mathscr{U}, R_1^-). \tag{2.17}$$

This identity can be verified by writing $\mathscr{G} = \sigma(\mathscr{U}, R_1^-, (R_i^-)_{i \geqslant 2})$ and integrating over generating rectangles, using (2.16) to express integrals in terms of conditional expectations given $\mathscr{U}$. Now to establish (2.15) we must show

$$E(\phi_1(R_1^+)\, \phi_2(R_2^+) | \mathscr{G}) = \prod E(\phi_i(R_i^+) | \mathscr{U}, R_i^-)$$

since by (2.17) the right side is $\prod E(\phi_i(R_i^+) | \mathscr{G})$. This identity is verified in the same manner as (2.17), writing $\mathscr{G} = \sigma(\mathscr{U}, R_1^-, R_2^-, (R_i^-)_{i \geqslant 3})$.

(b) Note first that $\sigma(A) = \sigma(R_i^- : i \leqslant 0) \subset \mathscr{U}$, for $\mathscr{U}$ defined as in (2.16). Applying Corollary 2.15(b) to the sequence $(R_{2i+1}, R_{2i+2})_{-\infty < i < \infty}$, we see that $E(\psi(R_1, R_2) | \mathscr{U})$ is essentially $\bigcap_{n > -\infty} \sigma(\psi(R_{2i+1}, R_{2i+2}) : i \leqslant n)$-measurable, for any ψ. Specializing to the case where $\psi(R_1, R_2)$ is of the form $\phi(R_1^-, R_2^-)$, we see that $E(\phi(R_1^-, R_2^-) | \mathscr{U})$ is essentially $\sigma(A)$-measurable, so $E(\phi(R_1^-, R_2^-) | \mathscr{U}) = E(\phi(R_1^-, R_2^-) | A)$. Since $\mathscr{U} = \sigma(A, C_j^- : j \geqslant 1)$, Lemma 2.2 shows that $\sigma(R_1^-, R_2^-)$ and $\sigma(C_j^- : j \geqslant 1)$ are c.i. given A. The same argument, replacing (R_1^-, R_2^-) by a k-vector $(R_1^-, ..., R_k^-)$, gives

$$\{\sigma(R_i^- : i \geqslant 1), \sigma(C_j^- : j \geqslant 1)\} \qquad \text{are c.i. given } A.$$

On the other hand, (2.16) applied to (R_i^-) shows that $\{R_i^- : i \geqslant 1\}$ are c.i. given A. Similarly $\{C_j^- : j \geqslant 1\}$ are c.i. given A. Hence, $\{C_1^-, C_2^-,...; R_1^-, R_2^-,...\}$ are c.i. given A, and (b) follows.

The proof of (c) requires two distinct applications of the following fact about bivariate exchangeable sequences.

LEMMA 2.18. *Let* $(Z_i)_{-\infty<i<\infty}$ *be exchangeable, where* $Z_i = (V_i, W_i)$. *Put* $\alpha^* = \sigma(Z_i: i \leqslant 0)$. *Fix* $j \geqslant 1$. *Then* $\{V_j, \sigma(\alpha^*, W_i: i \geqslant 1)\}$ *are c.i. given* $\sigma(\alpha^*, W_j)$.

Proof. We must verify

$$E(\phi(V_j)|\alpha^*, W_i: i \geqslant 1) = E(\phi(V_j)|\alpha^*, W_j). \tag{2.19}$$

We know from (2.6) that Z_j and $\sigma(Z_i: 1 \leqslant i \neq j)$ are c.i. given α^*, and this specializes to

$$\sigma(V_j, W_j) \quad \text{and} \quad \sigma(W_i: 1 \leqslant i \neq j) \qquad \text{are c.i. given } \alpha^*. \tag{2.20}$$

Use (2.20) to verify that the right-hand side of (2.19) integrates correctly over generating rectangles of $\sigma(\alpha^*, W_j, \sigma(W_i: 1 \leqslant i \neq j))$.

Proof of (c). Fix $k \geqslant 1$. Lemma 2.18, applied to $Z_i = (X_{k,i}, C_i^-)$, asserts

$$X_{k,j} \quad \text{and} \quad \sigma(A, R_k^-, C_i^- : i \geqslant 1) \qquad \text{are c.i. given} \quad \sigma(A, R_k^-, C_j^-). \tag{2.21}$$

Lemma 2.18 applied to $Z_i = (R_i^+, R_i^-)$ asserts that, for each $k \geqslant 1$,

$$R_k^+ \quad \text{and} \quad \mathscr{G} \qquad \text{are c.i. given} \quad \sigma(A, R_k^-, C_i^- : i \geqslant 1). \tag{2.22}$$

Now (c) is immediate because

$$E(\phi(X_{k,j})|\mathscr{G}) = E(\phi(X_{k,j})|A, R_k^-, C_i^- : i \geqslant 1) \qquad \text{by (2.22)}$$

since $X_{k,j}$ is a function of R_k^+,

$$= E(\phi(X_{k,j})|A, R_k^-, C_j^-) \qquad \text{by (2.21)}.$$

3. Further Structure of RCE Arrays

In this section we show how special forms of the general representation

$$X^*_{i,j} = f(\alpha, \xi_i, \eta_j, \lambda_{i,j}), \qquad i,j \geqslant 1, \tag{3.1}$$

correspond to structural properties of the RCE array $\mathbf{X}$.

By analogy with Silverman [14], call $\mathbf{X}$ *dissociated* if $(X_{i,j}: 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n)$ is independent of $(X_{i,j}: m < i, n < j)$ for each m, n. If $\mathbf{X}$ has a representation of the special form

$$X^*_{i,j} = g(\xi_i, \eta_j, \lambda_{i,j}), \tag{3.2}$$

then plainly $\mathbf{X}^*$ is dissociated.

PROPOSITION 3.3. *Let* $\mathbf{X}$ *be a dissociated RCE array. Then* $X \sim X^*$ *for some* $\mathbf{X}^*$ *of the form* (3.2).

Proof. The construction in the proof of Theorem 1.4 gave an array $\hat{\mathbf{X}} = (\hat{X}_{i,j})_{i,j \geqslant 1}$ which was shown to satisfy $\hat{\mathbf{X}} \sim \mathbf{X}$. The argument for this equivalence establishes also the slightly stronger equivalence $(A, \hat{\mathbf{X}}) \sim (A, \mathbf{X})$, where $A = (X_{i,j}: i,j \leqslant 0)$. But if $\mathbf{X}$ is dissociated, then $(X_{i,j}: i,j \geqslant 1)$ is independent of A. Hence, the final step in the proof of Theorem 1.4 (coding A as a function of α) gives a representation (3.1) in which $\mathbf{X}^*$ is independent of α. But then $X^*_{i,j} = f(a, \xi_i, \eta_j, \lambda_{i,j})$ for almost any $a \in [0, 1]$.

Remarks. The concept of disassociation arose from the following considerations, which we shall not develop in detail. It can be shown (by mimicking usual proofs of de Finetti's theorem, or by conditioning on α in (3.1)) that

each RCE array is a mixture of dissociated RCE arrays. (3.4)

From the viewpoint of Choquet theory, the dissociated arrays are precisely the extreme points of the class of all RCE arrays. At first sight, one might regard (3.3) rather than Theorem 1.4 as the "right" analogue of de Finetti's theorem for RCE arrays. But (3.3) is unsatisfactory because the structure of dissociated RCE arrays (unlike i.i.d. sequences) is not manifestly simple.

It is natural to ask when the representation (3.2) can be simplified to

$$X_{i,j} = h(\xi_i, \eta_j). \tag{3.5}$$

This question (essentially posed by Hoeffding [11]) is answered by Proposition 3.6 below, whose proof occupies the rest of this section. We introduce the *shell* σ-field $\mathscr{S} = \bigcap_{n \geqslant 1} \sigma(X_{i,j}: \max(i,j) \geqslant n)$.

PROPOSITION 3.6. *For a RCE array* $\mathbf{X}$ *the following are equivalent.*

(i) $\mathbf{X} \sim \mathbf{X}^*$ *for some* $\mathbf{X}^*$ *of the form* (3.4).

(ii) $\mathbf{X}$ *is dissociated and essentially* $\mathscr{S}$*-measurable.*

The proof is based on the lemma below. Throughout, let $\mathbf{X}$ be a dissociated array of the form (3.2), and put $\mathscr{F} = \sigma(\xi_i, \eta_j : i,j \geqslant 1)$.

LEMMA 3.7. (a) $\mathscr{S} \subset \mathscr{F}$ *a.s.*

(b) $\mathbf{X}$ *and* $\mathscr{F}$ *are c.i. given* $\mathscr{S}$.

Proof of Proposition 3.6. Suppose $X_{i,j} = h(\xi_i, \eta_j)$. Then $\mathbf{X}$ is $\mathscr{F}$-measurable. By Lemma 3.7(b), $\mathbf{X}$ and $\mathbf{X}$ are c.i. given $\mathscr{S}$. This easily implies $\mathbf{X}$ is essentially $\mathscr{S}$-measurable. Conversely, suppose $\mathbf{X}$ satisfies (ii). Without loss of generality we may suppose $X_{i,j}$ is uniformly bounded since we could replace $X_{i,j}$ by $\psi(X_{i,j})$ for some homeomorphism $\psi: \mathbb{R} \to (0,1)$. Now $E(X_{i,j} | \mathscr{F}) = h(\xi_i, \eta_j)$, where $h(x,y) = Eg(x,y,\lambda_{1,1})$. But $\mathbf{X}$ is essentially $\mathscr{S}$-measurable, so by Lemma 3.6(a) it is esentially $\mathscr{F}$-measurable, and so $X_{i,j} = E(X_{i,j} | \mathscr{F}) = h(\xi_i, \eta_j)$ a.s.

Proof of Lemma 3.7. To prove (a) put $\Lambda_n = \sigma(\lambda_{i,j} : \max(i,j) \geqslant n)$. Then $\mathscr{S} \subset \bigcap_n \sigma(\mathscr{F}, \Lambda_n)$. Now (Λ_n) is decreasing, $\bigcap_n \Lambda_n$ is trivial, and Λ_1 is independent of $\mathscr{F}$. These properties imply that $\bigcap_n \sigma(\mathscr{F}, \Lambda_n) = \mathscr{F}$ a.s.

For (b) we shall need the following curious result.

LEMMA 3.8. *Let H be a Hilbert space with inner product $\langle\ ,\ \rangle$. Let $(Y_i)_{i \geqslant 1}$ be i.i.d. H-valued random variables. Let $C_{i,j} = \langle Y_i, Y_j \rangle$ and let $\mathscr{C} = \sigma(C_{i,j} : 1 \leqslant i < j < \infty)$. Then there exist strictly increasing $\mathscr{C}$-measurable random indices $N_k \geqslant 2$ such that $Y_{N_k} \to Y_1$ a.s.*

Proof of Lemma 3.7(b). Suppose $X_{i,j} = f(\xi_i, \eta_j, \lambda_{i,j})$ takes values $\{0,1\}$ only. We first aim to prove

$$E(X_{1,1} | \mathscr{F}) = E(X_{1,1} | \mathscr{S}). \tag{3.9}$$

Define $p(x,y) = P(f(x,y,\lambda_{i,j}) = 1)$. Then $P(X_{1,1} = 1 | \mathscr{F}) = p(\xi_1, \eta_1)$, and so (3.9) is equivalent to the assertion

$$p(\xi_1, \eta_1) \qquad \text{is essentially } \mathscr{S}\text{-measurable.} \tag{3.10}$$

Define $c(x_1, x_2) = Ep(x_1, \eta_k) p(x_2, \eta_k)$. Fix (i,j) with $i \neq j$. By conditioning on (ξ_i, ξ_j) and applying the strong law of large numbers,

$$n^{-1} \sum_{k=1}^{n} X_{ik} X_{jk} \to c(\xi_i, \xi_j) \quad \text{a.s.}$$

Writing $\mathscr{F}_1 = \sigma(\xi_i : i \geqslant 1)$ we see that $c(\xi_i, \xi_j)$ is essentially $\mathscr{S} \cap \mathscr{F}_1$-measurable. Now for each $x \in [0,1]$ let $\hat{x}$ denote the function $p(x, \cdot)$ considered as an element of $H = L^2[0,1]$. The inner product $\langle \hat{x}_1, \hat{x}_2 \rangle$ is

$c(x_1, x_2)$. Apply Lemma 3.8 to $(\hat{\xi}_i)$. Then there exist $\mathscr{S} \cap \mathscr{F}_1$-measurable (N_k) such that, for almost all ω,

$$p(\xi_{N_k}(\omega), \cdot) \to p(\xi_1(\omega), \cdot) \qquad \text{in} \quad L^2[0, 1].$$

Similarly, writing $\mathscr{F}_2 = \sigma(\eta_j : j \geqslant 1)$, there exist $\mathscr{S} \cap \mathscr{F}_2$-measurable (M_k) such that, for almost all ω,

$$p(\cdot, \eta_{M_k}(\omega)) \to p(\cdot, \eta_1(\omega)) \qquad \text{in} \quad L^2[0, 1].$$

But $\mathscr{F}_1$ is independent of $\mathscr{F}_2$, so $(\xi_{N_k})_{k \geqslant 1}$ is independent of $(\eta_{M_k})_{k \geqslant 1}$. It easily follows that

$$p(\xi_{N_k}, \eta_{M_k}) \to p(\xi_1, \eta_1) \qquad \text{in} \quad L^2(\Omega).$$

Replacing (N_k, M_k) by a subsequence, we may assume the convergence is a.s. But $p(\xi_{N_k}, \eta_{M_k}) = P(X_{N_k, M_k} = 1 \mid \mathscr{F})$. So by conditioning on $\mathscr{F}$ and applying a law of large numbers,

$$n^{-1} \sum_{k=1}^{n} X_{N_k, M_k} \to p(\xi_1, \eta_1) \quad \text{a.s.}$$

This establishes (3.10), and hence (3.9).

Now consider a general array $X_{i,j} = f(\xi_i, \eta_j, \lambda_{i,j})$. Let ϕ be a measurable $\{0, 1\}$-valued function. Applying (3.9) to $\phi(X_{i,j})$ gives

$$E(\phi(X_{1,1}) \mid \mathscr{F}) = E(\phi(X_{1,1}) \mid \mathscr{S}_\phi),$$

where $\mathscr{S}_\phi$ is the shell σ-field of $(\phi(X_{i,j}))$. But $\mathscr{S}_\phi \subset \mathscr{S} \subset \mathscr{F}$ a.s., so

$$E(\phi(X_{1,1}) \mid \mathscr{F}) = E(\phi(X_{1,1}) \mid \mathscr{S}).$$

Standard arguments extend this to all bounded measurable ϕ. So

$$X_{1,1} \quad \text{and} \quad \mathscr{F} \qquad \text{are c.i. given } \mathscr{S}. \tag{3.11}$$

We finish with a trick. Fix $n \geqslant 2$ and define

$$\begin{aligned}
\xi_i^* &= (\xi_k : (i-1)\, n < k \leqslant in),\\
\eta_i^* &= (\eta_k : (i-1)\, n < k \leqslant in),\\
\lambda_{i,j}^* &= (\lambda_{k,m} : (i-1)\, n < k \leqslant in, (j-1)\, n < m \leqslant jn),\\
X_{i,j}^* &= (X_{k,m} : (i-1)\, n < k \leqslant in, (j-1)\, n < m \leqslant jn).
\end{aligned}$$

Then $X^*_{i,j} = f^*(\xi^*_i, \eta^*_j, \lambda^*_{i,j})$ for a certain function f^*. So (3.11) shows that $X^*_{1,1}$ and $\mathscr{F}^*$ are c.i. given $\mathscr{S}^*$. But $\mathscr{F}^* = \mathscr{F}$ and $\mathscr{S}^* = \mathscr{S}$. Letting $n \to \infty$ establishes Lemma 3.7(b).

Proof of Lemma 3.8. Write $\|\ \|$ for the norm in H. It is sufficient to prove

$$\|Y_1\| \quad \text{is } \mathscr{C}\text{-measurable} \tag{3.12}$$

for the same argument shows each $\|Y_j\|$ is $\mathscr{C}$-measurable, and hence so is $\|Y_j - Y_1\|^2 = \|Y_j\|^2 + \|Y_1\|^2 - 2\langle Y_1, Y_j\rangle$. Thus, we can define $N_k = \min\{j > N_{k-1}: \|Y_j - Y_1\| < 2^{-k}\}$, proving the lemma.

Let $(y_i)_{i \geqslant 1}$ be points in H such that

(a) the empirical distribution μ_n of $\{y_1, ..., y_n\}$ converges weakly, to μ_∞ say:

(b) each y_i lies in the support of μ_∞.

Write $c_{i,j} = \langle y_i, y_j\rangle$ for $i \neq j$. We shall show that $\|y_1\|$ can be defined in terms of the set $\{c_{i,j}\}$: applying this argument to $(Y_i(\omega))_{i \geqslant 1}$ proves (3.12).

Fix $a > b \geqslant 0$, $\delta > 0$. Call (a, b, δ) *permissible* if for $n \geqslant 1$ there exist $A_n \subset \{1, ..., n\}$ such that

(i) $1 \in A_n$,

(ii) $\liminf n^{-1} |A_n| \geqslant \delta$,

(iii) $a^2 \geqslant c_{i,j} \geqslant b^2$ for each $i, j \in A_n$.

For $\varepsilon > 0$ define $\alpha(\varepsilon) = \sup\{b: (b + \varepsilon, b, \delta)$ permissible for some $\delta > 0\}$. Then $\alpha(\varepsilon)$ decreases as ε decreases. We shall prove

$$\|y_1\| = \lim_{\varepsilon \to 0} \alpha(\varepsilon).$$

Let $G = \{x: \|x - y_1\| < \varepsilon/2\}$. By (b), $\mu_\infty(G) > 0$. By (a), $\liminf \mu_n(G) \geqslant \mu_\infty(G)$. Putting $A_n = \{i \leqslant n: y_i \in G\}$, we see that $(\|y_1\| + \varepsilon, \|y_1\| - \varepsilon, \mu_\infty(G))$ is permissible. Hence $\alpha(2\varepsilon) \geqslant \|y_1\| - \varepsilon$, and so $\|y_1\| \leqslant \lim_{\varepsilon \to 0} \alpha(\varepsilon)$. To obtain the reverse inequality, suppose (a, b, δ) is permissible. Let ν_n be the empirical distribution of $\{y_i: i \in A_n\}$. Then (ν_n) is tight, so some subsequence ν_{j_n} converges weakly to ν_∞ say. Consider a point z in the support of ν_∞. There must exist k_n, $K_n \in A_{j_n}$ such that $y_{k_n} \to z$, $y_{K_n} \to z$, $k_n \neq K_n$ for large n. It follows from (iii) that $\langle y_1, z\rangle \geqslant b^2$, $\langle z, z\rangle \leqslant a^2$. But $\langle y_1, z\rangle \leqslant \|y_1\| \|z\|$, and so we find that $\|y_1\| \geqslant b^2/a$, and the desired inequality follows easily.

Finally, we record two corollaries of Lemma 3.7 for use in the next section.

COROLLARY 3.13. *Suppose the elements of* $\mathbf{X}$ *are i.i.d. Then* $\mathbf{X}$ *is independent of* $\mathscr{F}$.

Proof. Here $\mathscr{S}$ is trivial.

COROLLARY 3.14. *Let* $\mathbf{X}$ *be dissociated. Then* $\{X_{i,j}: i,j \geqslant 1\}$ *are c.i. given* $\mathscr{S}$.

Proof. Representation (3.2) shows that $\{X_{i,j}: i,j \geqslant 1\}$ are c.i. given $\mathscr{F}$. Now apply Lemma 3.7.

4. SPHERICAL ARRAYS

Let $\mathbf{Y} = (Y_{i,j})_{i,j \geqslant 1}$ be an array,. and for each n consider the subarray $Y^n = (Y_{i,j}: 1 \leqslant i,j \leqslant n)$ as an $(n \times n)$ matrix. Following Dawid [1], call $\mathbf{Y}$ *spherical* if for each n

$$U_1 Y^n U_2 \sim Y^n \qquad \text{for all orthogonal } (n \times n) \text{ matrices } U_1, U_2. \tag{4.1}$$

Our aim is to characterize spherical arrays. Write $R_i^n = (Y_{i,1},..., Y_{i,n})$, and let $\mathscr{E}_n$ be the tail σ-field of $(R_i^n)_{i \geqslant 1}$. Call $\mathbf{Y}$ *row-presentable* if for each n there exists a random $(n \times n)$ matrix Σ_n such that, conditional on $\mathscr{E}_n$, the sequence $(R_1^n, R_2^n,...)$ is i.i.d. with distribution $N(0, \Sigma_n)$. Write $\mathbf{Y}^T$ for the transposed array $Y_{i,j}^T = Y_{j,i}$. Dawid [1, 2] established the following properties.

PROPOSITION 4.2. *Suppose* $\mathbf{Y}$ *is spherical. Then* $\mathbf{Y}$ *is row-presentable, and* $\mathbf{Y} \sim \mathbf{Y}^T$.

Dawid [2] also observes that a spherical array is a mixture of spherical arrays with independent diagonal entries: his argument can be extended to show that a spherical array is a mixture of dissociated spherical arrays. For this reason, we shall restrict attention to the dissociated case. As (4.1) immediately implies that a spherical array is RCE, we may employ representation (3.2) in analyzing a dissociated spherical array.

Call $\mathbf{Y}$ *Normal* if its entries $(Y_{i,j})$ are independent $N(0, 1)$. Call $\mathbf{Y}$ *product-Normal* if $\mathbf{Y} = (V_i \cdot W_j)_{i,j \geqslant 1}$, where $(V_1, V_2,...; W_1, W_2,...)$ are independent $N(0, 1)$. Dawid [2] observed that Normal and product-Normal arrays are spherical, and so are linear combinations of independent such arrays. He conjectured the general spherical array was a mixture of arrays produced in this manner. We shall prove this, under a moment condition.

THEOREM 4.3. *Let* $\mathbf{Y}$ *be a dissociated spherical array, and suppose* $E(Y_{1,1})^2 < \infty$. *Then* $\mathbf{Y} \sim a_0 \mathbf{Y}^0 + \sum a_n \mathbf{Y}^n$, *where* $\sum a_n^2 < \infty$, $\mathbf{Y}^0$ *is Normal,* $\mathbf{Y}^n$ $(n \geqslant 1)$ *is product-Normal, and* $(\mathbf{Y}^n)_{n \geqslant 0}$ *are independent.*

The proof is based on the lemmas below. As in Section 3, let $Y_{i,j} =$

$g(\xi_i, \eta_j, \lambda_{i,j})$ be the representation (3.2), $\mathscr{S}$ the shell σ-field, and $\mathscr{F} = \sigma(\xi_i, \eta_j: i, j \geq 1)$. Write $E(\mathbf{Y}|\mathscr{S})$ for the array $E(Y_{i,j}|\mathscr{S})$.

LEMMA 4.4. *$E(\mathbf{Y}|\mathscr{S})$ and $\mathbf{Y} - E(\mathbf{Y}|\mathscr{S})$ are spherical.*

LEMMA 4.5. *If $E(Y_{1,1}|\mathscr{S}) = 0$, then $\mathbf{Y} = a_0\mathbf{Z}$ for some Normal $\mathbf{Z}$.*

LEMMA 4.6. *If $\mathbf{Y}$ has the form $Y_{i,j} = h(\xi_i, \eta_j)$, then $\mathbf{Y} = \sum a_n \mathbf{Z}^n$, where $\sum a_n^2 < \infty$ and $(\mathbf{Z}^n)_{n \geq 1}$ are independent $\mathscr{F}$-measurable product-Normal arrays.*

Proof of Theorem 4.3. Define

$$\begin{aligned} Y^1_{i,j} &= h(\xi_i, \eta_j), \qquad \text{where} \quad h(x, y) = Eg(x, y, \lambda_{i,j}); \\ Y^2_{i,j} &= Y_{i,j} - Y^1_{i,j} \\ &= f(\xi_i, \eta_j, \lambda_{i,j}), \qquad \text{where} \quad f = g - h. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{Y}^1 &= E(\mathbf{Y}|\mathscr{F}) \\ &= E(\mathbf{Y}|\mathscr{S}) \qquad \text{by Lemma 3.7(b).} \end{aligned} \tag{4.7}$$

Thus, $\mathbf{Y}^1$ and $\mathbf{Y}^2$ are spherical by Lemma 4.4, and clearly each is dissociated. Lemma 4.6 gives an expression for $\mathbf{Y}^1$ in terms of F-measurable arrays. To handle $\mathbf{Y}^2$, note that its shell σ-field

$$\begin{aligned} \mathscr{S}^2 &= \bigcap \sigma(Y^2_{i,j}: \max(i,j) > n) \\ &= \bigcap \sigma(Y_{i,j} - E(Y_{i,j}|\mathscr{S}): \max(i,j) > n) \qquad \text{by (4.7)} \\ &\subset \bigcap \mathscr{S}_n = \mathscr{S}. \end{aligned}$$

But $E(\mathbf{Y}^2|\mathscr{S}) = E(\mathbf{Y} - \mathbf{Y}^1|\mathscr{S}) = 0$, and so $E(\mathbf{Y}^2|\mathscr{S}^2) = 0$. Now Lemma 4.5 shows $\mathbf{Y}^2$ is a multiple of a Normal array. Finally, Corollary 3.13 shows that $\mathbf{Y}^2$ is independent of $\mathscr{F}$.

Proof of Lemma 4.4. Fix $k < m < n$. Define $\mathscr{S}_{m,n} = \sigma(Y_{i,j}: m < \max(i,j) \leq n)$. Let U_1, U_2 be orthogonal $(k \times k)$ matrices, and write $\hat{U}_1$, $\hat{U}_2$ for the $(n \times n)$ matrices

$$\hat{U}_r = \begin{pmatrix} U_r & 0 & 0 \\ 0 & I_{[k+1,m]} & 0 \\ 0 & 0 & I_{[m+1,n]} \end{pmatrix}$$

Expanding in blocks, we can write the relation $Y_{[1,n]} \sim \hat{U}_1 Y_{[1,n]} \hat{U}_2$ in the form

$$\begin{pmatrix} Y_{[1,k]} & A & B \\ C & Y_{[k+1,m]} & D \\ F & G & Y_{[m+1,n]} \end{pmatrix} \sim \begin{pmatrix} U_1 Y_{[1,k]} U_2 & U_1 A & U_1 B \\ C U_2 & Y_{[k+1,m]} & D \\ F U_2 & G & Y_{[m+1,n]} \end{pmatrix}.$$

Now

$$\begin{aligned} \mathscr{S}_{m,n} &= \sigma(B, D, Y_{[m+1,n]}, F, G) \\ &= \sigma(U_1 B, D, Y_{[m+1,n]}, F U_2, G) \qquad \text{as } U_r \text{ is invertible.} \end{aligned}$$

So

$$(Y_{[1,k]}, E(Y_{[1,k]} | \mathscr{S}_{m,n})) \sim (U_1 Y_{[1,k]} U_2, E(U_1 Y_{[1,k]} U_2 | \mathscr{S}_{m,n})).$$

By the martingale convergence theorems we may let $n \to \infty$, then $m \to \infty$, obtaining

$$(Y_{[1,k]}, E(Y_{[1,k]} | \mathscr{S})) \sim (U_1 Y_{[1,k]} U_2, U_1 E(Y_{[1,k]} | \mathscr{S})\, U_2)$$

and the lemma follows.

Proof of Lemma 4.5. Let $\mathscr{E}_2$ denote the tail σ-field of the sequence $(Y_{i,1}, Y_{i,2})_{i \geqslant 1}$. Proposition 4.2 shows that, conditional on $\mathscr{E}_2$, the random vectors $(Y_{i,1}, Y_{i,2})$ are independent $N(0, \Sigma_2)$. We want to get information about the random matrix Σ_2.

Clearly $\mathscr{E}_2 \subset \mathscr{S}$. By hypothesis $E(Y_{i,j} | \mathscr{S}) = 0$, so by Corollary 3.14 $E(Y_{1,1} Y_{1,2} | \mathscr{S}) = 0$. Thus $E(Y_{1,1} Y_{1,2} | \mathscr{E}_2) = 0$. In other words, conditional on $\mathscr{E}_2$ the pair $Y_{1,1}$, $Y_{1,2}$ are orthogonal. Hence, Σ_2 must be of the form

$$\begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

for some random σ_1 and σ_2. By the strong law of large numbers,

$$n^{-1} \sum_{i=1}^{n} Y_{2i,1}^2 \to \sigma_1^2 \qquad \text{a.s.}$$

$$n^{-1} \sum_{i=1}^{n} Y_{2i+1,2}^2 \to \sigma_2^2 \qquad \text{a.s.,}$$

so by dissociativity σ_1 and σ_2 are independent. And $\sigma_1 \sim \sigma_2$ by column-exchangeability.

On the other hand, writing $\beta = E|V|$ for $V \sim N(0, 1)$.

$$E(|Y_{1,1}Y_{1,2}|\,|\mathscr{C}_2) = \beta^2\sigma_1\sigma_2,$$
$$E(|Y_{1,1}Y_{2,1}|\,|\mathscr{C}_2) = \beta^2\sigma_1^2.$$

Hence

$$\begin{aligned} E\sigma_1^2 &= \beta^{-2}E\,|Y_{1,1}Y_{1,2}| \\ &= \beta^{-2}E\,|Y_{1,1}Y_{2,1}| && \text{by Proposition 4.2} \\ &= E\sigma_1\sigma_2 \\ &= (E\sigma_1)^2 && \text{by independence.} \end{aligned}$$

So $\sigma_1 = \sigma_2 = a_0$ a.s., for some constant a_0. The same argument holds for Σ_k, $k \geqslant 2$, and establishes the lemma.

Proof of Lemma 4.6. Write (ξ, η) for (ξ_1, η_1). Because $h(\xi, \eta)$ is square-integrable, a theorem in classical analysis (see Smithies [15, Theorem 8.3.3]) gives a biorthogonal expansion

$$h(\xi, \eta) = \sum a_k \theta_k(\xi)\, \psi_k(\eta), \tag{4.8}$$

where $\sum a_k^2 < \infty$, each of the families $(\theta_k(\xi))_{k \geqslant 1}$ and $(\psi_k(\eta))_{k \geqslant 1}$ is orthonormal, and the sum converges in L^2. We shall prove that $(\psi_k(\eta))$ are multivariate-Normal: a similar argument holds for the θ's, and establishes the lemma.

For $x \in [0, 1]$ let $\hat{x}$ denote the random variable $h(x, \eta)$. Suppose forst that for each $(x_1, \dots, x_n)$,

$$(\hat{x}_1, \dots, \hat{x}_n) \qquad \text{is multivariate-Normal.} \tag{4.9}$$

Let H be the Hilbert space of sequences $(y_i)_{i \geqslant 1}$ with $\sum y_i^2 < \infty$. Let H_0 be the subset of H for which $\sum y_i \psi_i(\eta)$ converges in L^2 to some Normal limit. Then H_0 is closed in H: we must prove $H_0 = H$. By (4.8) there exists a subset A of $[0, 1]$ such that $\xi \in A$ a.s. and

$$\sum a_k \theta_k(x)\, \psi_k(\eta) \to \hat{x} \quad \text{in} \quad L^2, \qquad x \in A. \tag{4.10}$$

Let H_1 be the subspace of H consisting of finite linear combinations of $(a_i \theta_i(x))_{i \geqslant 1}$, $x \in A$. Using (4.10) and (4.9), we see that $H_1 \subset H_0$. Now suppose $(y_i) \in H$ is orthogonal to H_1. Then, $\sum a_i y_i \theta_i(x) = 0$, $x \in A$. So $\sum a_i y_i \theta_i(\xi) = 0$ a.s., and because $(\theta_i(\xi))$ are orthogonal we must have $y_i = 0$. Hence $\bar{H}_1 = H$, and so $H_0 = H$ as desired.

It remains to prove (4.9). Fix n, and put $R_i = (Y_{i,1},..., Y_{i,n})$. Recall $Y_{i,j} = h(\xi_i, \eta_j)$. So as $k \to \infty$ the empirical distribution of $(R_1(\omega),..., R_k(\omega))$ converges, for almost all ω, to the distribution $(\hat{\xi}_1(\omega),..., \hat{\xi}_n(\omega))$. But from Proposition 4.2 the empirical distribution of $(R_1(\omega), R_2(\omega),..., R_k(\omega))$ converges, for almost all ω, to $N(0, \Sigma_n)$. Thus, (4.9) certainly holds for almost all $(x_1,..., x_n) \in [0, 1]^n$. Now consider the map $x \to \hat{x}$ as a random variable $\Gamma\colon [0, 1] \to L^2[0, 1]$. Let $B = \{x\colon \hat{x} \in \text{support}(\Gamma)\}$. Then B has measure 1, and it is easily checked that $(\hat{x}_1,..., \hat{x}_n)$ is multivariate-Normal if $x_i \in B$ for each i. Replacing $h(x, y)$ by $h(x, y)1_{(x \in B)}$, we may assume (4.9) holds everywhere.

5. Related Results

(a) *Partial exchangeability.* Given a family $\mathbf{Z} = (Z_i)_{i \in I}$ and a subgroup Π_0 of the group of permutations of I, it is natural to call $\mathbf{Z}$ *partially exchangeable* (with respect to Π_0) if (1. 1) holds for $\pi \in \Pi_0$. Clearly RCE arrays are precisely the families indexed by $I = \mathbb{N} \times \mathbb{N}$ which are partially exchangeable with respect to the direct product of the permutation groups. It is natural to ask whether there are characterizations analogous to Theorem 1.4 for other partially exchangeable families. This question seems difficult, and we know only one other result. Hoeffding [10] introduced U-statistics, which can be regarded as normalized partial sums of arrays of the form $X_{\{i,j\}} = g(\xi_i, \xi_j) + g(\xi_j, \xi_i)$. Such arrays are partially exchangeable with respect to permutations $\{i, j\} \to \{\pi(i), \pi(j)\}$. Silverman [14] and Eagleson and Weber [8] call this property *weak exchangeability*, and generalize limit theorems for U-statistics to weakly exchangeable arrays. Here is a characterization: the proof is omitted, being similar to the proof of Theorem 1.4.

THEOREM 5.1. *Given a weakly exchangeable array* $\mathbf{X}$, *there exists a measurable function* $f(a, b, c, d)$ *with* $f(a, \cdot, \cdot, d)$ *symmetric for fixed* (a, d), *such that* $\mathbf{X} \sim \mathbf{X}^* = (f(a, \xi_i, \xi_j, \lambda_{\{i,j\}}))$.

(b) *Second-order structure.* A forthcoming monograph of Terry Speed *et al.* is devoted to a profound investigation of analysis of variance from the viewpoint of second-order partial exchangeability (where only the covariance structure is assumed to be invariant).

(c) *Bayesian statistics.* Kingman [13] discusses briefly the possible interpretation of Theorem 1.4 in the context of Bayesian statistics.

(d) *Visual perception.* A $\{0, 1\}$-valued RCE array may be viewed as a random pattern of black and white squares. Essentially different arrays seem to yield visually distinguishable patterns, and this has consequences for hypotheses about visual perception—see Diaconis and Freedman [5].

(e) *Other mixtures.* Another type of generalization of de Finetti's theorem is to characterize mixtures of processes other than i.i.d. sequences. For work on mixtures of Markov processes see Freedman [9] and Diaconis and Freedman [6].

Acknowledgments

The ideas in the present proof of Theorem 1.4 come from John Kingman, and are a great simplification of the author's original proof. Phil Dawid suggested Theorem 1.4 could be used to prove his conjecture, Theorem 4.3. The inspiration for using orthonormal families in Lemma 4.6 came from Geoff Eagleson's work on U-statistics. Research carried out in 1978–1979 during the tenure of a Research Fellowship at St. John's College, Cambridge.

References

[1] Dawid, A. P. (1977). Spherical matrix distributions and a multivariate model. *J. Roy. Statist. Soc. Ser. B* **39** 254–261.

[2] Dawid, A. P. (1978). Extendibility of spherical matrix distributions. *J. Multivariate Anal.* **8** 567–572.

[3] Dellacherie, C. and Meyer, P.-A. (1975). *Probabilités et Potentiel*, Chaps. I–IV. Hermann, Paris.

[4] Dellacherie, C. and Meyer, P.-A. (1980). *Probabilités et potentiel*, Chaps. V–VII. Hermann, Paris.

[5] Diaconis, P. and Freedman, D. (1981). The statistics of the Julesz conjecture in visual perception. *J. Math. Psych.* in press.

[6] Diaconis, P. and Freedman, D. (1980). De Finetti's theorem for Markov chains. *Ann. Probab.* **8** 115–130.

[7] Dynkin, E. B. (1978). Sufficient statistics and extreme points. *Ann. Probab.* **6** 705–730.

[8] Eagleson, G. K. and Weber, N. C. (1979). Limit theorems for weakly exchangeable arrays. *Proc. Cambridge Philos. Soc.* **84** 73–80.

[9] Freedman, D. (1962). Invariants under mixing which generalize de Finetti's theorem. *Ann. Math. Statist.* **33** 916–923.

[10] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[11] Hoeffding, W. (1977). De Finetti's theorem and U-statistics. Preprint.

[12] Kingman, J. F. C. (1978). Uses of exchangeability. *Ann. Probab.* **6** 183–197.

[13] Kingman, J. F. C. (1979). Contribution to discussion on: The reconciliation of probability assessments. *J. Roy. Statist. Soc. Ser. A* **142** 171.

[14] Silverman, B. W. (1976). Limit theorems for dissociated random variables. *Adv. in Appl. Probab.* **8** 806–819.

[15] Smithies, F. (1958). *Integral Equations*. Cambridge Univ. Press, London/New York.

[16] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.