ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 31, Number 3, Fall 2001

A RESTRICTED DICHOTOMY OF EQUIVALENCE CLASSES FOR SOME MEASURES OF DEPENDENCE

RICHARD C. BRADLEY

ABSTRACT. In some limit theory for weakly dependent random sequences, a role is implicitly played by certain measures of dependence based on "covariances" of random variables taking their values in Banach spaces. Here it is shown that in a certain restricted sense, there is a "dichotomy of equivalence classes" for measures of dependence of that type that involve "∞-norms" of the random variables. The question of a possible corresponding "unrestricted" dichotomy of equivalence classes remains open.

1. Introduction. In probability theory, there is a large literature on limit theorems under "strong mixing conditions." The formulations of such mixing conditions are based on "measures of dependence" between σ-fields of events. Some of that limit theory involves random variables taking their values in a Hilbert space or (more generally) in a Banach space.

Building on the work of Rosenblatt [**13**, Chapter 7], Dehling and Philipp [**7**] and other researchers, the author, Bryc, and Janson wrote a series of papers [**4**], [**5**], [**6**] on the relationships (e.g., "dominations" or "equivalencies") within certain classes of measures of dependence. The latter paper [**6**] studied in detail a broad class of measures of dependence involving "covariances" of random variables taking their values in general Hilbert spaces or Banach spaces.

For the measures of dependence of that latter type that involve the "∞-norms" of those random variables, it turns out that there is what one might refer to as a "restricted dichotomy of equivalence classes." It will be formulated in Theorem 1.8 and Remark 1.12 below. The question of a possible corresponding "unrestricted dichotomy" remains open; more on that in Remark 1.12.

This work was partially supported by NSF grant DMS-9703712.
Received by the editors on May 3, 2000, and in revised form on July 1, 2000.

For simplicity, the presentation is given here for real Banach spaces. However, the results hold as well for complex Banach spaces, e.g., by a trivial extension of Theorem 1.8.

Definition 1.1. Suppose $(\Omega, \mathcal{M}, P)$ is a probability space. For any two σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, define the following three "measures of dependence":

$$\alpha(\mathcal{F}, \mathcal{G}) := \sup_{\substack{F \in \mathcal{F} \\ G \in \mathcal{G}}} |P(F \cap G) - P(F)P(G)|; \tag{1.1}$$

$$\psi^*(\mathcal{F}, \mathcal{G}) := \sup_{\substack{F \in \mathcal{F} \\ G \in \mathcal{G} \\ P(F)P(G) > 0}} \frac{P(F \cap G)}{P(F)P(G)}; \tag{1.2}$$

$$\beta(\mathcal{F}, \mathcal{G}) := \sup \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |P(F_i \cap G_j) - P(F_i)P(G_j)| \tag{1.3}$$

where this latter supremum is taken over all pairs of (finite) partitions $\{F_1, F_2, \dots, F_I\}$ and $\{G_1, G_2, \dots, G_J\}$ of Ω such that $F_i \in \mathcal{F}$ for each i and $G_j \in \mathcal{G}$ for each j.

Remark 1.2. In (1.3), the factor of $1/2$ is not particularly important. It has become customary in the literature on strong mixing conditions. In (1.2), the notation $\psi^*(.,.)$ is used because the notation $\psi(.,.)$ has a well-established different (though closely related) meaning. One has that

$$0 \le 2\alpha(\mathcal{F}, \mathcal{G}) \le \beta(\mathcal{F}, \mathcal{G}) \le 1 \le \psi^*(\mathcal{F}, \mathcal{G}) \le \infty. \tag{1.4}$$

The first and third inequalities in (1.4) are trivial consequences of (1.1) and (1.3). The fourth inequality in (1.4) holds by (1.2) and the fact that if $P(F \cap G) < P(F) \cdot P(G)$, then $P(F^c \cap G) > P(F^c) \cdot P(G)$. The second inequality in (1.4) holds by a simple calculation and the fact that for events F and G, the quantity $|P(F \cap G) - P(F)P(G)|$ remains unchanged if F is replaced by its complement F^c and/or G is replaced by G^c. The fifth inequality in (1.4) is trivial; the point of it is that the equality $\psi^*(\mathcal{F}, \mathcal{G}) = \infty$ can occur.

Remark 1.3. The measures of dependence $\alpha(.,.)$ and $\beta(.,.)$ are the bases for, respectively, the "α-mixing" (or "strong mixing") condition introduced by Rosenblatt [**12**], and the "β-mixing" (or "absolute regularity") condition introduced by Volkonskii and Rozanov [**15**] (and attributed there to Kolmogorov). (For strictly stationary, finite-state random sequences, β-mixing is equivalent to the "weak Bernoulli" condition of Ornstein isomorphism theory; see, e.g., Shields [**14**]). As a consequence of (1.4), β-mixing implies α-mixing; however, α-mixing does not imply β-mixing. The formulations of these two mixing conditions need not be given here.

For the moment, let H be an arbitrary separable real Hilbert space, and let B be an arbitrary separable real Banach space. An extremely sharp central limit theorem proved by Doukhan, Massart and Rio [**8**, Theorem 1] for real-valued random variables under α-mixing, was extended to H-valued random variables under α-mixing by Merlevède, Peligrad and Utev [**10**, Theorem 1.3]. Dehling and Philipp [**7**, Theorem 1] proved a very sharp almost sure invariance principle for H-valued random variables under α-mixing. Dehling and Philipp [**7**, Theorem 4] also proved a very sharp almost sure invariance principle for B-valued random variables under β-mixing; it is not known whether that result holds (for B-valued random variables) under α-mixing. The author [**3**] established a connection between tightness of sums and tightness of linear functionals of sums, for B-valued random variables under β-mixing; it is not known whether (for B-valued random variables) that result still holds under α-mixing. For limit theory for H-valued random variables, α-mixing seems to be a natural mixing condition; and for limit theory for B-valued random variables, β-mixing seems to be natural. (One exception to this general pattern is a result of Philipp [**11**, Theorem 2] on convergence of normalized sums of B-valued random variables to stable laws under α-mixing, an extension of an earlier corresponding result in Ibragimov and Linnik [**9**, Theorem 18.1.1] for real-valued random variables under α-mixing.)

Definition 1.4. Suppose B is a (not necessarily separable) nontrivial real Banach space, with norm $\|\cdot\|_B$ and B^* is its (real) dual space, the Banach space of real bounded linear functionals on B, with norm $\|\cdot\|_{B^*}$. That is, for $y \in B^*$, $\|y\|_{B^*} := \sup_{x\in B} |\langle x, y\rangle|/\|x\|_B$ (where $0/0 := 0$). Here and below, we use the notation $\langle x, y\rangle := y(x)$ for $x \in B$ and $y \in B^*$. (This notation "fits" the well-known connection between linear functionals and "inner products" in the case of a real Hilbert space.)

Recall our probability space $(\Omega, \mathcal{M}, P)$. For any two σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, define the following measure of dependence:

$$R(B; \mathcal{F}, \mathcal{G}) := \sup |E\langle X, Y\rangle - \langle EX, EY\rangle| \tag{1.5}$$

where the supremum is taken over all pairs of *simple* random variables X and Y such that X is B-valued and $\mathcal{F}$-measurable, Y is B^*-valued and $\mathcal{G}$-measurable, and $\|X\|_B \le 1$ a.s. and $\|Y\|_{B^*} \le 1$ a.s. (By using only simple random variables here, one avoids certain measure-theoretic technicalities.)

Remark 1.5. Suppose $\mathcal{F}$ and $\mathcal{G}$ are any σ-fields $\subset \mathcal{M}$.

One has that

$$4\alpha(\mathcal{F}, \mathcal{G}) = R(\mathbf{R}; \mathcal{F}, \mathcal{G}) \le R(B; \mathcal{F}, \mathcal{G}) \le 2\beta(\mathcal{F}, \mathcal{G}). \tag{1.6}$$

These facts are elementary. By [**9**, Theorem 17.2.1], one has that $R(\mathbf{R}; \mathcal{F}, \mathcal{G}) \le 4\alpha(\mathcal{F}, \mathcal{G})$. To see that equality holds there, consider the (real-valued) random variables of the form $X = I(F) - I(F^c)$, $F \in \mathcal{F}$ and $Y = I(G) - I(G^c)$, $G \in \mathcal{G}$. (Here $I(.)$ denotes the indicator function.) Also, $R(\mathbf{R}; \mathcal{F}, \mathcal{G}) \le R(B; \mathcal{F}, \mathcal{G})$ by a simple argument involving an application of the Hahn-Banach theorem. The last inequality in (1.6) will take slightly more work to verify.

Suppose X, respectively Y, is a simple $\mathcal{F}$-measurable, respectively $\mathcal{G}$-measurable, random variable taking its values in the unit ball of B, respectively B^*. Represent $X = \sum_{i=1}^{I} x_i I(F_i)$ and $Y = \sum_{j=1}^{J} y_j I(G_j)$ where $\{F_1, \dots, F_I\}$, respectively $\{G_1, \dots, G_J\}$, is a partition of Ω into events in $\mathcal{F}$, respectively $\mathcal{G}$, and $x_i \in B$, $\|x_i\|_B \le 1$, $y_j \in B^*$, and $\|y_j\|_{B^*} \le 1$. Then, by a simple calculation,

$$\begin{aligned} &E\langle X, Y\rangle - \langle EX, EY\rangle \\ &\qquad = \sum_{i=1}^{I}\sum_{j=1}^{J} \langle x_i, y_j\rangle [P(F_i \cap G_j) - P(F_i)P(G_j)]. \end{aligned} \tag{1.7}$$

Now $|\langle x_i, y_j \rangle| \le 1$ for each (i, j), and hence the last inequality in (1.6) follows from (1.5) and (1.3).

Using Grothendieck's inequality and the first equality in (1.6), Dehling and Philipp [**7**, Lemma 2.2] pointed out that for any real (not necessarily separable) Hilbert space H,

$$R(H; \mathcal{F}, \mathcal{G}) \le 10\alpha(\mathcal{F}, \mathcal{G}). \tag{1.8}$$

On the other hand, one has that

$$R(c_0; \mathcal{F}, \mathcal{G}) = R(l^\infty; \mathcal{F}, \mathcal{G}) = R(l^1; \mathcal{F}, \mathcal{G}) = 2\beta(\mathcal{F}, \mathcal{G}). \tag{1.9}$$

The second and third equalities were pointed out in [**6**, Theorem 3.1] with a simple argument, and that argument yields the first equality as well. Here c_0, l^1 and l^∞ refer to the usual Banach spaces of sequences of real numbers (with c_0 being the subspace of l^∞ consisting of the sequences that converge to 0).

In a sense, equations (1.8) and (1.9) "mirror" the observations in Remark 1.3. For more on this, see [**6**].

We shall return to Banach spaces in Remark 1.12 below, after dealing with a broader class of measures of dependence. The next remark will help provide a framework for what follows.

Remark 1.6. For every $\delta > 0$, there exist a probability space $(\Omega, \mathcal{M}, P)$ and a pair of σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$ such that

$$\alpha(\mathcal{F}, \mathcal{G}) \le \delta, \qquad \beta(\mathcal{F}, \mathcal{G}) = 1/2, \quad \text{and } \psi^*(\mathcal{F}, \mathcal{G}) = 2. \tag{1.10}$$

One can see this from the class of examples presented in [**6**, pp. 431–433]. (The dependence coefficient $\alpha(\mathcal{F}, \mathcal{G})$ can be fit into the analysis there via the first equality in (1.6) above. The property $\psi^*(\mathcal{F}, \mathcal{G}) = 2$ was not mentioned there, but is easy to verify for those examples.)

In what follows, for simplicity, we shall use that class of examples as a "benchmark," and restrict our stated upper bounds on $\psi^*(\mathcal{F}, \mathcal{G})$ to numbers η, $2 \le \eta \le \infty$.

Definition 1.7. Suppose Θ and Γ are nonempty sets and $\chi : \Theta \times \Gamma \to [-1, 1]$ is a function. For any given probability space $(\Omega, \mathcal{M}, P)$ and any two σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, define the following "measure of dependence":

(1.11)
$$R_{\Theta,\Gamma,\chi}(\mathcal{F},\mathcal{G}) := \sup \Big| \sum_{i=1}^{I} \sum_{j=1}^{J} \chi(\theta_i,\gamma_j) \cdot [P(F_i \cap G_j) - P(F_i)P(G_j)] \Big|$$

where this supremum is taken over all pairs of partitions $\{F_1, \dots, F_I\}$ and $\{G_1, \dots, G_J\}$ of Ω with $F_i \in \mathcal{F}$ for each i and $G_j \in \mathcal{G}$ for each j and all choices of (not necessarily distinct) elements $\theta_1, \dots, \theta_I \in \Theta$ and $\gamma_1, \dots, \gamma_J \in \Gamma$.

Theorem 1.8. *Suppose $2 \le \eta < \infty$ and $0 < C < 1$. Then there exists a positive number $\tau := \tau(\eta, C)$ such that the following holds*:

Suppose Θ and Γ are nonempty sets, and $\chi : \Theta \times \Gamma \to [-1,1]$ is a function such that

$$\forall\, \theta \in \Theta,\ \exists \tilde{\theta} \in \Theta \text{ such that } \forall\, \gamma \in \Gamma, \quad \chi(\tilde{\theta},\gamma) = -\chi(\theta,\gamma). \tag{1.12}$$

Suppose that for every $\delta > 0$, there exist a probability space $(\Omega, \mathcal{M}, P)$ and a pair of σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$ such that

$$\alpha(\mathcal{F},\mathcal{G}) \le \delta, \qquad R_{\Theta,\Gamma,\chi}(\mathcal{F},\mathcal{G}) > C, \quad \text{and } \psi^*(\mathcal{F},\mathcal{G}) \le \eta. \tag{1.13}$$

Then for every probability space $(\Omega, \mathcal{M}, P)$ and every pair of σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, one has that

$$R_{\Theta,\Gamma,\chi}(\mathcal{F},\mathcal{G}) \ge \tau \cdot \beta(\mathcal{F},\mathcal{G}). \tag{1.14}$$

Remark 1.9. Theorem 1.8 will be proved in Section 2.

It is unknown whether Theorem 1.8 still holds for $\eta = \infty$ (making the dependence coefficient $\psi^*(.,.)$ irrelevant). In the proof (in Section 2) the use of the bound $\psi^*(\mathcal{F},\mathcal{G}) \le \eta$ seems to be crucial in the derivation of equations (2.16.16)–(2.16.17).

Obviously there is no essential change in Theorem 1.8 if one restricts to finite σ-fields. Theorem 1.8 can be reformulated in terms of a certain class of matrices (such as matrices with the entries p_{ij} in (2.5.1) or the entries p^*_{mn} in (2.18.1) in Section 2).

With a change in the positive number τ, Theorem 1.8 can be adapted to functions χ that map $\Theta \times \Gamma$ into some other reasonable set besides $[-1, 1]$, such as a bounded set in the complex plane.

The following definition will facilitate further discussions.

Definition 1.10. Suppose Θ and Γ are nonempty sets, and $\chi : \Theta \times \Gamma \to [-1, 1]$ is a function.

The ordered triplet (Θ, Γ, χ) is said to satisfy "Condition $\mathcal{B}$" if there exists a positive number $\tau = \tau(\Theta, \Gamma, \chi)$ such that for every probability space $(\Omega, \mathcal{M}, P)$ and every pair of σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$, equation (1.14) holds.

If $2 \le \eta \le \infty$, then the ordered triplet (Θ, Γ, χ) is said to satisfy "Condition $\mathcal{A}(\eta)$" if the following holds: For every $\varepsilon > 0$, there exists $\delta = \delta(\eta, \varepsilon, \Theta, \Gamma, \chi) > 0$, such that for every probability space $(\Omega, \mathcal{M}, P)$ and every pair of σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$ such that $\psi^*(\mathcal{F}, \mathcal{G}) \le \eta$ and $\alpha(\mathcal{F}, \mathcal{G}) \le \delta$ hold, one has that $R_{\Theta,\Gamma,\chi}(\mathcal{F}, \mathcal{G}) \le \varepsilon$.

Obviously, by (1.3) and (1.11), one has that $R_{\Theta,\Gamma,\chi}(\mathcal{F}, \mathcal{G}) \le 2\beta(\mathcal{F}, \mathcal{G})$. Condition $\mathcal{B}$ says that (for the given Θ, Γ, χ), the measures of dependence $R_{\Theta,\Gamma,\chi}(.,.)$ and $\beta(.,.)$ are within a positive constant factor of each other.

For a given η, $2 \le \eta \le \infty$, condition $\mathcal{A}(\eta)$ says (for a given Θ, Γ, χ) that under the restriction $\psi^*(.,.) \le \eta$, a "small" value of $\alpha(.,.)$ forces a "small" value of $R_{\Theta,\Gamma,\chi}(.,.)$.

Obviously, if $2 \le \eta_1 < \eta_2 \le \infty$, then condition $\mathcal{A}(\eta_2)$ implies condition $\mathcal{A}(\eta_1)$. Of course, in condition $\mathcal{A}(\infty)$, the dependence coefficient $\psi^*(.,.)$ is irrelevant.

Corollary 1.11. *Suppose* Θ *and* Γ *are nonempty sets and* $\chi : \Theta \times \Gamma \to [-1, 1]$ *is a function such that* (1.12) *holds. Then the following three statements hold*:

(1) *For the ordered triplet* (Θ, Γ, χ)*, the conditions* $\mathcal{A}(\eta)$*,* $2 \le \eta < \infty$*, are equivalent.*

(2) *For the ordered triplet* (Θ, Γ, χ), *exactly one of the following two statements* (a), (b) *holds*:

(a) *Condition* $\mathcal{A}(\eta)$ *is satisfied for all* η, $2 \le \eta < \infty$.

(b) *Condition* $\mathcal{B}$ *is satisfied.*

(3) *Suppose that for every* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon, \Theta, \Gamma, \chi) > 0$ *such that for every probability space* $(\Omega, \mathcal{M}, P)$ *and every pair of* σ-*fields* $\mathcal{F}$ *and* $\mathcal{G} \subset \mathcal{M}$ *such that* $R_{\theta,\Gamma,\chi}(\mathcal{F}, \mathcal{G}) \le \delta$ *holds, one has that* $\beta(\mathcal{F}, \mathcal{G}) \le \varepsilon$. *Then condition* $\mathcal{B}$ *holds.*

Proof. Here we take Theorem 1.8 for granted.

Proof of (1). Suppose that for some $\eta \in [2, \infty)$, condition $\mathcal{A}(\eta)$ fails to hold. Then by Theorem 1.8, condition $\mathcal{B}$ holds. Applying this to the examples cited in Remark 1.6, one has that condition $\mathcal{A}(2)$ fails to hold. Hence (trivially) for every $\eta \in [2, \infty)$, condition $\mathcal{A}(\eta)$ fails to hold. Part (1) follows.

Proof of (2). If statement (a) in part (2) fails to hold, then statement (b) there holds by Theorem 1.8. Conversely if statement (b) holds, then applying that to the examples cited in Remark 1.6, one has that statement (a) cannot hold. Part (2) follows.

Proof of (3). If the hypothesis of (3) holds, then by the examples cited in Remark 1.6, condition $\mathcal{A}(2)$ fails to hold, and hence by (say) part (2), condition $\mathcal{B}$ holds. Part (3) follows.

Remark 1.12. Now let us return to Banach spaces. Suppose B is a (not necessarily separable) nontrivial real Banach space. Let Θ, respectively Γ, denote the unit ball of B, respectively of B^*, and define $\chi : \Theta \times \Gamma \to [-1, 1]$ by $\chi(x, y) := \langle x, y \rangle$. Then $R(B; \mathcal{F}, \mathcal{G}) = R_{\Theta,\Gamma,\chi}(\mathcal{F}, \mathcal{G})$ by (1.5), (1.7) and (1.11). We shall say that B satisfies condition $\mathcal{B}$, respectively condition $\mathcal{A}(\eta)$ for a given η, $2 \le \eta \le \infty$, if this ordered triplet (Θ, Γ, χ) satisfies condition $\mathcal{B}$, respectively condition $\mathcal{A}(\eta)$. Note that equation (1.12) is satisfied, since $\langle -x, y \rangle = -\langle x, y \rangle$.

By (1.8), (real) Hilbert spaces satisfy condition $\mathcal{A}(\infty)$. Certain other (real) Banach spaces seem to be known (at least in principle) to satisfy condition $\mathcal{A}(\infty)$, by results in interpolation theory on Banach spaces (see, e.g., Bergh and Löfström [**2**, Chapter 5] or Bennett and Sharpley [**1**]). By (1.9), the Banach spaces c_0, l^1 and l^∞ satisfy condition $\mathcal{B}$.

By Corollary 1.11, for a given (real) Banach space B, the conditions $\mathcal{A}(\eta)$, $2 \le \eta < \infty$, are equivalent; and either B satisfies the conditions $\mathcal{A}(\eta)$, $2 \le \eta < \infty$, or B satisfies condition $\mathcal{B}$, but not both. (Similarly, Corollary 1.11(3) applies to this context.) The following open question remains unsolved: If B satisfies conditions $\mathcal{A}(\eta)$, $2 \le \eta < \infty$, does it follow that B satisfies condition $\mathcal{A}(\infty)$? If the answer is affirmative, then B would satisfy either condition $\mathcal{A}(\infty)$ or condition $\mathcal{B}$.

Here is another perspective. As in [**4**], [**5**], [**6**], let us say that two measures of dependence are "equivalent" if each one becomes arbitrarily small as the other becomes sufficiently small. By (1.6) and (1.8), for any nontrivial (real) separable Hilbert space H, the measure of dependence $R(H; ., .)$ is equivalent to $\alpha(., .)$. That is, for arbitrary pairs of σ-fields $\mathcal{F}$ and $\mathcal{G}$ on an arbitrary probability space $(\Omega, \mathcal{M}, P)$, one has that $R(H; \mathcal{F}, \mathcal{G})$ and $\alpha(\mathcal{F}, \mathcal{G})$ each become arbitrarily small as the other becomes sufficiently small. By (1.9), the measures of dependence $R(c_0; ., .), R(l^1; ., .), R(l^\infty; ., .)$ and $\beta(., .)$ are equivalent. By Remark 1.6, the measures of dependence $\alpha(., .)$ and $\beta(., .)$ are not equivalent. It remains an open question whether for every nontrivial (real) Banach space B, the measure of dependence $R(B; \dots)$ is equivalent to one of the two measures of dependence $\alpha(., .)$ or $\beta(., .)$. If so, that would be an ("unrestricted") "dichotomy of equivalence classes" for Banach spaces B. By (1.6) and Corollary 1.11 (parts (1) and (2)), one at least has the following: For every nontrivial (real) Banach space B, either (i) the measure of dependence $R(B; ., .)$ is equivalent to $\beta(., .)$ or (ii) there is a "restricted" equivalence of $R(B; ., .)$ with $\alpha(., .)$, with the pairs of σ-fields $\mathcal{F}$ and $\mathcal{G}$ being limited to ones satisfying a fixed finite upper bound on $\psi^*(\mathcal{F}, \mathcal{G})$. This is the "restricted dichotomy of equivalence classes" alluded to earlier.

2. Proof of Theorem 1.8. The rest of this paper is devoted to the proof of Theorem 1.8. Throughout this proof, the following conventions will be used:

(i) The (one-dimensional) Lebesgue measure of a Borel set $S \subset \mathbf{R}$ will be denoted [Leb. meas. S].

(ii) The cardinality of a finite set S will be denoted [card. S].

(iii) An "empty sum" $\sum_{i\in\varnothing} c_i$ (where $\varnothing$ is the empty set) is defined to be zero.

(iv) For typographical convenience, a sum of the form $\sum_{i\in S} c_i$ will sometimes be written as $\sum[c_i \mid i \in S]$ and a sum of the form $\sum_{(i,j)\in S} c_{ij}$ will sometimes be written $\sum[c_{ij} \mid (i,j) \in S]$. (The summand and the set of indices will be separated by a vertical line.)

In order for the proof to be easier to follow, it will be broken into a sequence of small "steps" (including some definitions and lemmas). The numbering of equations will be based on those steps.

Step 2.1. As in the hypothesis of Theorem 1.8, suppose

$$2 \le \eta < \infty \quad \text{and} \quad 0 < C < 1. \tag{2.1.1}$$

Our first task is to define the positive number $\tau := \tau(\eta, C)$.

Define the positive numbers A_k, $1 \le k \le 8$ and τ as follows

$$A_1 := C/10, \tag{2.1.2}$$
$$A_2 := A_1/(3\eta), \tag{2.1.3}$$
$$A_3 := A_2/2, \tag{2.1.4}$$
$$A_4 := A_3/3, \tag{2.1.5}$$
$$A_5 := A_2A_4/2, \tag{2.1.6}$$
$$A_6 := A_5C/40, \tag{2.1.7}$$
$$A_7 := \max((0, A_6C/60] \cap \{1/2, 1/3, 1/4, 1/5, \dots\}), \tag{2.1.8}$$
$$A_8 := A_3A_7, \tag{2.1.9}$$

and

$$\tau := \max((0, A_8/2] \cap \{1/2, 1/3, 1/4, 1/5, \dots\}). \tag{2.1.10}$$

This completes the definition of the positive number $\tau = \tau(\eta, C)$. Note that by (2.1.1)–(2.1.10),

$$A_1, A_2, \dots, A_8, \tau \in (0, 1). \tag{2.1.11}$$

Also, in (2.1.8) and (2.1.10), the main requirements are $0 < A_7 \le A_6C/60$ and $0 < \tau \le A_8/2$. Having A_7 and τ be reciprocals of positive integers is not vital, but will be helpful for "bookkeeping" purposes later on.

Step 2.2. Now (as in the statement of Theorem 1.8) suppose that Θ and Γ are nonempty sets, and $\chi : \Theta \times \Gamma \to [-1, 1]$ is a function such that (1.12) holds:

$$\forall\, \theta \in \Theta,\ \exists \tilde{\theta} \in \Theta \text{ such that } \forall\, \gamma \in \Gamma, \quad \chi(\tilde{\theta}, \gamma) = -\chi(\theta, \gamma). \tag{2.2.1}$$

Also, suppose that for every $\delta > 0$, there exist a probability space and a pair of σ-fields on that space such that (1.13) holds.

To complete the proof of Theorem 1.8, our task is to prove that for every probability space and every pair of σ-fields on that space, (1.14) holds with the positive number τ defined in (2.1.10).

Let $(\Omega^*, \mathcal{M}^*, P^*)$ be an arbitrary fixed probability space, and let $\mathcal{F}^*$ and $\mathcal{G}^*$ be arbitrary fixed σ-fields $\subset \mathcal{M}^*$. In order to complete the proof of Theorem 1.8, it suffices to prove that

$$R_{\Theta,\Gamma,\chi}(\mathcal{F}^*, \mathcal{G}^*) \geq \tau \cdot \beta(\mathcal{F}^*, \mathcal{G}^*). \tag{2.2.2}$$

The rest of Section 2 is devoted to the proof of this inequality.

Step 2.3. Refer to (1.3) and the last paragraph of Step 2.2. Let M and N be positive integers, and $\{F_1^*, F_2^*, \ldots, F_M^*\}$ and $\{G_1^*, G_2^*, \ldots, G_N^*\}$ be partitions of Ω^*, with $F_m^* \in \mathcal{F}^*$ for each m and $G_n^* \in \mathcal{G}^*$ for each n, such that

$$\sum_{m=1}^{M} \sum_{n=1}^{N} |P^*(F_m^* \cap G_n^*) - P^*(F_m^*)P^*(G_n^*)| \geq \beta(\mathcal{F}^*, \mathcal{G}^*). \tag{2.3.1}$$

We need to define some more positive constants. Now $A_2 > A_3$ by (2.1.4) (and (2.1.11)). Let B_1 be a number in $(0, 1)$ such that

$$A_2 - 2B_1 > A_3. \tag{2.3.2}$$

Also, $A_3 > A_4$ by (2.1.5) (and (2.1.11)). Referring to (2.1.1), let B_2 be a number in $(0, 1)$ such that

$$A_3 - A_4 - 2\eta B_2 > 0. \tag{2.3.3}$$

Referring to (2.1.11), define the number B_3 in $(0,1)$ by

$$B_3 := (\tau/2) \cdot B_2. \tag{2.3.4}$$

Recall the positive integer M above (from (2.3.1)). Referring to (2.1.1), define the number B_4 in $(0,1)$ by

$$B_4 := B_3^M C/40. \tag{2.3.5}$$

Referring to (2.3.3), let δ_0 be a number in $(0,1)$ such that

$$\frac{\delta_0}{(A_3 - A_4 - 2\eta B_2) \cdot B_4} < B_1. \tag{2.3.6}$$

For later reference, note again that

$$B_1, B_2, B_3, B_4, \delta_0 \in (0,1). \tag{2.3.7}$$

In order to prove (2.2.2) and thereby complete the proof of Theorem 1.8, we need to select some key elements from the sets Θ and Γ to use in conjunction with Definition 1.7 and the partitions $\{F_1^*, \dots, F_M^*\}$ and $\{G_1^*, \dots, G_N^*\}$ of Ω^*. The task of selecting those key elements from Θ and Γ will involve three stages: First, in Steps 2.4–2.17, we shall do extensive preliminary work on another, separate probability space. Next, in Steps 2.18–2.23, we shall return to the probability space $(\Omega^*, \mathcal{M}^*, P^*)$ and do some more preliminary work. Finally, in Step 2.24, the key elements from Θ and Γ will be selected and (2.2.2) will be proved (and thereby the proof of Theorem 1.8 will be complete).

Step 2.4. Refer to the statement of Theorem 1.8 and the second paragraph of Step 2.2. Let $(\Omega, \mathcal{M}, P)$ be a probability space and $\mathcal{F}$ and $\mathcal{G}$ be σ-fields $\subset \mathcal{M}$ such that (1.13) holds with $\delta = \delta_0$:

$$\alpha(\mathcal{F}, \mathcal{G}) \le \delta_0; \quad R_{\Theta,\Gamma,\chi}(\mathcal{F}, \mathcal{G}) > C; \quad \text{and} \quad \psi^*(\mathcal{F}, \mathcal{G}) \le \eta. \tag{2.4.1}$$

(Here, of course, δ_0, C and η are as in (2.1.1) and (2.3.6).)

Referring to (2.4.1) and Definition 1.7, let I and J be positive integers and $\{F_1, \dots, F_I\}$ and $\{G_1, \dots, G_J\}$ be partitions of Ω with $F_i \in \mathcal{F}$ for each i and $G_j \in \mathcal{G}$ for each j, and $\theta_1, \dots, \theta_I \in \Theta$ and $\gamma_1, \dots, \gamma_J \in \Gamma$ such that

$$\Big| \sum_{i=1}^{I} \sum_{j=1}^{J} \chi(\theta_i, \gamma_j) \cdot [P(F_i \cap G_j) - P(F_i)P(G_j)] \Big| > C.$$

Refer to (2.2.1) (a restatement of (1.12)). Replacing θ_i by $\tilde{\theta}_i$ for each i if necessary, we assume that

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \chi(\theta_i, \gamma_j) \cdot [P(F_i \cap G_j) - P(F_i)P(G_j)] > C. \tag{2.4.2}$$

Step 2.5. Recall the probability space $(\Omega, \mathcal{M}, P)$, the integers I and J, and the events F_i and G_j from Step 2.4. For each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$, define the number

$$p_{ij} := P(F_i \cap G_j). \tag{2.5.1}$$

For each $i \in \{1, \dots, I\}$, define the number

$$a_i := P(F_i) = \sum_{j=1}^{J} p_{ij}. \tag{2.5.2}$$

(The second equality follows trivially from (2.5.1).) For each $j \in \{1, \dots, J\}$, define the number

$$b_j := P(G_j) = \sum_{i=1}^{I} p_{ij}. \tag{2.5.3}$$

For each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$, define the number

$$\varepsilon_{ij} := P(F_i \cap G_j) - P(F_i)P(G_j) = p_{ij} - a_i b_j. \tag{2.5.4}$$

The following observations are elementary. First,

$$p_{ij} \geq 0 \text{ for each } (i, j); \quad \text{and} \quad \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} = 1. \tag{2.5.5}$$

Also,

$$a_i \geq 0 \text{ for each } i; \quad b_j \geq 0 \text{ for each } j; \qquad (2.5.6)$$
$$\sum_{i=1}^{I} a_i = 1; \quad \text{and} \quad \sum_{j=1}^{J} b_j = 1.$$

Also,

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \varepsilon_{ij} = 0. \qquad (2.5.7)$$

Also, for each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$,

$$-a_i b_j \leq \varepsilon_{ij} \leq p_{ij} \leq \eta a_i b_j. \qquad (2.5.8)$$

Here the last inequality holds because by (2.4.1) (even if $P(F_i) = 0$ or $P(G_j) = 0$),

$$p_{ij} = P(F_i \cap G_j) \leq \psi^*(\mathcal{F}, \mathcal{G}) \cdot P(F_i)P(G_j) \leq \eta a_i b_j.$$

Next, suppose Q and S are any sets such that $Q \subset \{1, \dots, I\}$ and $S \subset \{1, \dots, J\}$. Consider the events $F := \cup_{i \in Q} F_i$ and $G := \cup_{j \in S} G_j$. By (1.1) and (2.4.1), $|P(F \cap G) - P(F)P(G)| \leq \delta_0$. Of course, $P(F) = \sum_{i \in Q} a_i$ and $P(G) = \sum_{j \in S} b_j$, and also $P(F \cap G) = \sum_{i \in Q} \sum_{j \in S} p_{ij}$. Thus, by (2.5.4), one has the following:

$$\forall\, Q \subset \{1, \dots, I\}, \quad \forall\, S \subset \{1, \dots, J\}, \qquad (2.5.9)$$
$$\Big| \sum_{i \in Q} \sum_{j \in S} \varepsilon_{ij} \Big| = \Big| \sum_{i \in Q} \sum_{j \in S} (p_{ij} - a_i b_j) \Big| \leq \delta_0.$$

Step 2.6. For each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$, define the number

$$r_{ij} := (1/2)(p_{ij} + a_i b_j). \qquad (2.6.1)$$

By (2.5.1), (2.5.2), (2.5.3), (2.5.5) and (2.5.6), these numbers have the following properties:

$$r_{ij} \geq 0 \text{ for all } (i,j); \quad \text{and} \quad \sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij} = 1; \tag{2.6.2}$$

$$\forall\, i \in \{1, \dots, I\}, \quad \sum_{j=1}^{J} r_{ij} = a_i; \tag{2.6.3}$$

and

$$\forall\, j \in \{1, \dots, J\}, \quad \sum_{i=1}^{I} r_{ij} = b_j. \tag{2.6.4}$$

Also, by (2.5.4) and (2.6.1), for each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$,

$$|\varepsilon_{ij}| \leq 2r_{ij}. \tag{2.6.5}$$

Next, recall from Step 2.4 the elements $\theta_1, \dots, \theta_I \in \Theta$ and $\gamma_1, \dots, \gamma_J \in \Gamma$ such that (2.4.2) holds. For each $i \in \{1, \dots, I\}$ and each $j \in \{1, \dots, J\}$, define the notation

$$f_{ij} := \chi(\theta_i, \gamma_j). \tag{2.6.6}$$

Then by (2.4.2), (2.5.4) and the assumption (in Theorem 1.8) that χ maps $\Theta \times \Gamma$ into $[-1, 1]$, one has that

$$-1 \leq f_{ij} \leq 1 \quad \text{for all } (i,j); \tag{2.6.7}$$

and

$$\sum_{i=1}^{I}\sum_{j=1}^{J} f_{ij}\varepsilon_{ij} > C. \tag{2.6.8}$$

Lemma 2.7. *There exists a Borel set $E_0 \subset [-1, 1]$ with the following two properties*:

(i) [Leb. meas. E_0] $\geq 2 - C/5$,

and

(ii) *for every* $x \in E_0$, *one has that*

$$\sum [r_{ij} | (i,j) : x - A_7 < f_{ij} < x + A_7] \le A_6. \tag{2.7.1}$$

Here, of course, A_6 and A_7 are the positive numbers from (2.1.7) and (2.1.8). Of course, the sum in (2.7.1) is

$$\sum [r_{ij} \mid (i,j) \in \{1,\dots,I\} \times \{1,\dots,J\} : x - A_7 < f_{ij} < x + A_7],$$

the sum of r_{ij}, taken over all $(i,j) \in \{1,\dots,I\} \times \{1,\dots,J\}$ such that $x - A_7 < f_{ij} < x + A_7$. Throughout the rest of this paper, it will be tacitly understood that i, respectively j, always means an element of $\{1,\dots,I\}$, respectively $\{1,\dots,J\}$, and that the "set of all (i,j) such that ..." means the "set of all $(i,j) \in \{1,\dots,I\} \times \{1,\dots,J\}$ such that ... ."

Proof. Referring to (2.1.8), define the integer $\kappa := 1/A_7$. Of course $\kappa \ge 2$. Define the closed interval $\mathcal{I}_1 := [-1, -1 + 2A_7]$. For each $k = 2, 3, \dots, \kappa$, define the half-open interval $\mathcal{I}_k := (-1 + 2(k-1)A_7, -1 + 2kA_7]$. These intervals $\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_\kappa$ are (pairwise) disjoint, and their union is $[-1, 1]$.

For each $k = 1, 2, \dots, \kappa$, define the nonnegative number (see (2.6.2))

$$\mu_k := \sum [r_{ij} \mid (i,j) : f_{ij} \in \mathcal{I}_k]. \tag{2.7.2}$$

By (2.6.2) and (2.6.7), $\sum_{k=1}^{\kappa} \mu_k = \sum_{i=1}^{I} \sum_{j=1}^{J} r_{ij} = 1$.

Let $T(1)$ denote the set of all $k \in \{1,\dots,\kappa\}$ such that $\mu_k \ge A_6/2$. Then

$$1 \ge \sum_{k \in T(1)} \mu_k \ge \sum_{k \in T(1)} A_6/2 = (A_6/2) \cdot [\text{card.}\, T(1)].$$

Hence, $[\text{card.}\, T(1)] \le 2/A_6$. Let $T(2)$ denote the set of all $k \in \{1,\dots,\kappa\}$ such that the set $\{k-1, k, k+1\} \cap T(1)$ is nonempty. Then $[\text{card.}\, T(2)] \le 3 \cdot [\text{card.}\, T(1)] \le 6/A_6$.

Define the set $E_1 := \cup_{k \in T(2)} \mathcal{I}_k$. Now each of the intervals $\mathcal{I}_k$ has Lebesgue measure $2A_7$. Hence, by (2.1.8),

$$[\text{Leb. meas. } E_1] = 2A_7 \cdot [\text{card.}\, T(2)] \leq (2A_7) \cdot (6/A_6) \leq C/5.$$

Define the set $E_0 := [-1, 1] - E_1$. Then [Leb. meas. E_0] $\geq 2 - C/5$. This gives property (i) in Lemma 2.7.

Now let x be an arbitrary, fixed element of E_0. To complete the proof of property (ii) in Lemma 2.7, it suffices to prove (2.7.1) for this x.

The open interval $(x - A_7, x + A_7)$ contains at most one of the points -1, $-1 + 2A_7$, $-1 + 4A_7, \dots, -1 + 2\kappa A_7$ $(= 1)$. Hence, there exists an element $k \in \{1, 2, \dots, \kappa - 1\}$ such that $(x - A_7, x + A_7) \cap [-1, 1] \subset \mathcal{I}_k \cup \mathcal{I}_{k+1}$. Neither k nor $k+1$ is an element of the set $T(1)$. (Otherwise, one would have that k and $k+1$ are both in $T(2)$ and $x \in \mathcal{I}_k \cup \mathcal{I}_{k+1} \subset E_1$, contradicting the stipulation that $x \in E_0$.) It follows from (2.7.2) and (2.6.7) (and (2.6.2) and the definition of $T(1)$) that

$$\begin{aligned} \sum [r_{ij} \,|\, (i, j) : x - A_7 < f_{ij} < x + A_7] & \\ \leq \mu_k + \mu_{k+1} < (A_6/2) + (A_6/2) = A_6. & \end{aligned}$$

Thus (2.7.1) holds. This completes the proof of property (ii) in Lemma 2.7.

Step 2.8. Refer to (2.6.7). Let L be the positive integer and $g_0, g_1, \dots, g_L$ be the numbers in $[-1, 1]$ such that

$$\{g_0, g_1, \dots, g_L\} = \{-1, 1\} \cup \{f_{ij} : 1 \leq i \leq I, 1 \leq j \leq J\}, \tag{2.8.1}$$

and

$$-1 = g_0 < g_1 < g_2 < \cdots < g_{L-1} < g_L = 1. \tag{2.8.2}$$

That is, the g_l's are, in increasing order, the numbers f_{ij} together with -1 and $+1$.

Also, define the function $\Phi : [-1, 1] \to \mathbf{R}$ as follows. For each $x \in [-1, 1]$,

$$\Phi(x) := \sum [(f_{ij} - x)\varepsilon_{ij} \,|\, (i, j) : f_{ij} \leq x]. \tag{2.8.3}$$

Lemma 2.9. *The function* Φ *in* (2.8.3) *has the following five properties*:

(i) $\Phi(-1) = 0$.

(ii) $\Phi(1) > C$.

(iii) Φ *is continuous on* $[-1, 1]$.

(iv) Φ *is differentiable on the set* $[-1, 1] - \{g_0, g_1, \dots, g_L\}$; *and for each* $l \in \{0, 1, \dots, L-1\}$ *and each* $x \in (g_l, g_{l+1})$, *one has that*

$$\Phi'(x) = -\sum[\varepsilon_{ij} \,|\, (i,j) : f_{ij} \le g_l] = -\sum[\varepsilon_{ij} \,|\, (i,j) : f_{ij} \le x]. \tag{2.9.1}$$

(v) $|\Phi'(x)| \le 2$ *for all* $x \in [-1, 1] - \{g_0, g_1, \dots, g_L\}$.

Proof. To verify property (i), note that by (2.6.7) and (2.8.3),

$$\Phi(-1) = \sum[(f_{ij} + 1)\varepsilon_{ij} \,|\, (i,j) : f_{ij} = -1] = 0.$$

Property (ii) holds since, by (2.6.7), (2.8.3), (2.5.7) and (2.6.8),

$$\Phi(1) = \sum_{i=1}^{I}\sum_{j=1}^{J}(f_{ij} - 1)\varepsilon_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{J} f_{ij}\varepsilon_{ij} > C.$$

Proof of property (iii). Refer to (2.8.1) and (2.8.2). It suffices to prove that Φ is continuous on each of the closed intervals $[g_l, g_{l+1}]$, $l \in \{0, 1, \dots, L-1\}$.

Let $l \in \{0, 1, \dots, L-1\}$ be arbitrary but fixed. For each $x \in [g_l, g_{l+1})$ (the point g_{l+1} is excluded for now), one has trivially by (2.8.1), (2.8.2) and (2.8.3) that

$$\Phi(x) = \sum[(f_{ij} - x)\varepsilon_{ij} \,|\, (i,j) : f_{ij} \in \{g_0, g_1, \dots, g_l\}]. \tag{2.9.2}$$

Hence Φ is continuous on the (half-open) interval $[g_l, g_{l+1})$. Also, by (2.9.2), with g_{l+1} also written as $g(l+1)$, for typographical convenience,

(2.9.3)
$$\lim_{x \to g(l+1)-} \Phi(x) = \sum[(f_{ij} - g_{l+1})\varepsilon_{ij} \,|\, (i,j) : f_{ij} \in \{g_0, g_1, \dots, g_l\}].$$

However, trivially,

$$\sum [(f_{ij} - g_{l+1})\varepsilon_{ij} \,\big|\, (i,j) : f_{ij} = g_{l+1}] = 0.$$

Hence by (2.8.3) (with (2.8.1) and (2.8.2)), the right side of (2.9.3) equals $\Phi(g_{l+1})$. That is, by (2.9.3), $\lim_{x \to g(l+1)-} \Phi(x) = \Phi(g_{l+1})$. Hence Φ is continuous on the closed interval $[g_l, g_{l+1}]$. Since $l \in \{0, 1, \ldots, L-1\}$ was arbitrary, property (iii) follows.

Proof of property (iv). Let $l \in \{0, 1, \ldots, L-1\}$ be arbitrary but fixed. To complete the proof of (iv), it suffices to show that (2.9.1) holds for all $x \in (g_l, g_{l+1})$.

Recall from the proof of property (iii) that (2.9.2) holds for $x \in (g_l, g_{l+1})$. Differentiating (2.9.2), one obtains the first equality in (2.9.1) (for $x \in (g_l, g_{l+1})$). The second equality in (2.9.1) follows (for $x \in (g_l, g_{l+1})$) from (2.8.1) and (2.8.2). This completes the proof of property (iv).

Property (v) follows from property (iv) since $\sum_{i=1}^{I} \sum_{j=1}^{J} |\varepsilon_{ij}| \le 2$ by (say) (2.6.2) and (2.6.5). This completes the proof of Lemma 2.9.

Lemma 2.10. *There exists a real number ζ with the following four properties*:

(2.10.1) $$-1 < \zeta < 1;$$

(2.10.2) $$\sum [r_{ij} \,\big|\, (i,j) : \zeta - A_7 < f_{ij} < \zeta + A_7] \le A_6;$$

(2.10.3) $$\sum [\varepsilon_{ij} \,\big|\, (i,j) : f_{ij} \le \zeta - A_7] \le -C/10;$$

and

(2.10.4) $$\sum [\varepsilon_{ij} \,\big|\, (i,j) : f_{ij} \ge \zeta + A_7] \ge C/10.$$

Proof. Referring to (2.8.1) and (2.8.2), define the set

(2.10.5) $$S(1) := [-1, 1] - \{g_0, g_1, \ldots, g_L\}.$$

Refer to (2.8.3). By Lemma 2.9, the function Φ is continuous on $[-1, 1]$ and its derivative Φ' is defined, bounded and continuous on $S(1)$. (In fact by (2.9.1), Φ' is constant on each of the open intervals (g_l, g_{l+1}), $l = 0, 1, \dots, L-1$.) Also, the set $\{g_0, g_1, \dots, g_L\}$, the complement of $S(1)$ in $[-1, 1]$, has only finitely many elements. It follows from Lemma 2.9 that

$$\int_{-1}^{1} \Phi'(x)\, dx = \Phi(1) - \Phi(-1) > C. \tag{2.10.6}$$

Define the sets

$$S(2) := \{x \in S(1) : \Phi'(x) \geq C/4\} \tag{2.10.7}$$

and

$$S(3) := \{x \in S(1) : \Phi'(x) < C/4\}. \tag{2.10.8}$$

By (2.10.8) (and (2.1.1)),

$$\int_{S(3)} \Phi'(x)\, dx \leq \int_{S(3)} (C/4)\, dx \leq \int_{-1}^{1} (C/4)\, dx = C/2.$$

Hence, by (2.10.6), (2.10.7) and (2.10.8), $\int_{S(2)} \Phi'(x)\, dx \geq C/2$. Hence by Lemma 2.9 (v), $\int_{S(2)} 2\, dx \geq C/2$. Hence

$$[\text{Leb. meas. } S(2)] \geq C/4. \tag{2.10.9}$$

Let the set E_0 be as in Lemma 2.7. Then from (2.10.9) and property (i) in Lemma 2.7, one has that

$$[\text{Leb. meas. } E_0] + [\text{Leb. meas. } S(2)] \geq 2 - (C/5) + C/4 > 2.$$

Since E_0 and $S(2)$ are each a subset of $[-1, 1]$, it follows that $E_0 \cap S(2)$ is nonempty. Let ζ be an element of $E_0 \cap S(2)$. Of course by (2.8.1), (2.10.5) and (2.10.7), neither 1 nor -1 is a member of $S(2)$. Hence equation (2.10.1) holds. Since $\zeta \in E_0$, equation (2.10.2) also holds (see property (ii) in Lemma 2.7). Now the remaining task is to prove (2.10.3)–(2.10.4).

Since $\zeta \in S(2)$, one has that $\Phi'(\zeta) \geq C/4$ by (2.10.7). Hence by Lemma 2.9 (iv),

$$\sum[\varepsilon_{ij} \,|\, (i,j) : f_{ij} \leq \zeta] \leq -C/4. \tag{2.10.10}$$

Also, by (2.6.5), (2.10.2), (2.1.7) and (2.1.11),

$$\begin{aligned}&\Big|\sum[\varepsilon_{ij} \,|\, (i,j) : \zeta - A_7 < f_{ij} \leq \zeta]\Big| \\ &\qquad \leq \sum[2r_{ij} \mid (i,j) : \zeta - A_7 < f_{ij} \leq \zeta] \leq 2A_6 \leq C/10.\end{aligned} \tag{2.10.11}$$

Hence by (2.10.10), equation (2.10.3) holds.

Next by (2.10.10) and (2.5.7),

$$\sum[\varepsilon_{ij} \,|\, (i,j) : f_{ij} > \zeta] \geq C/4. \tag{2.10.12}$$

By an argument analogous to (2.10.11),

$$\Big|\sum[\varepsilon_{ij} \,|\, (i,j) : \zeta < f_{ij} < \zeta + A_7]\Big| \leq C/10.$$

Hence by (2.10.12), equation (2.10.4) holds.

All four equations (2.10.1)–(2.10.4) in Lemma 2.10 have been verified. This completes the proof of Lemma 2.10.

Step 2.11. Henceforth, let ζ be a fixed number satisfying all properties in Lemma 2.10.

Referring to (2.6.6), define the numbers h_{ij} and λ_{ij}, $i \in \{1, \ldots, I\}$, $j \in \{1, \ldots, J\}$ as follows:

$$h_{ij} := f_{ij} - \zeta; \tag{2.11.1}$$

and

$$\lambda_{ij} := \begin{cases} 1 & \text{if } h_{ij} \geq A_7 \\ 0 & \text{if } -A_7 < h_{ij} < A_7 \\ -1 & \text{if } h_{ij} \leq -A_7. \end{cases} \tag{2.11.2}$$

For convenient reference, let us list, in this new terminology, the four properties in Lemma 2.10:

$$-1 < \zeta < 1; \tag{2.11.3}$$

$$\sum [r_{ij} \,|\, (i,j) : \lambda_{ij} = 0] \le A_6; \tag{2.11.4}$$

$$\sum [\varepsilon_{ij} \,|\, (i,j) : \lambda_{ij} = -1] \le -C/10; \tag{2.11.5}$$

$$\sum [\varepsilon_{ij} \,|\, (i,j) : \lambda_{ij} = 1] \ge C/10. \tag{2.11.6}$$

Also, by (2.6.7) and (2.11.1), one has that for each $i \in \{1, \dots, I\}$ and $j \in \{1, \dots, J\}$,

$$-1 - \zeta \le h_{ij} \le 1 - \zeta. \tag{2.11.7}$$

By (2.11.2), (2.11.5) and (2.11.6), one also has that

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \lambda_{ij} \varepsilon_{ij} \ge C/5. \tag{2.11.8}$$

Lemma 2.12. *Suppose S is a nonempty subset of $\{1, \dots, J\}$. Suppose i is an element of $\{1, \dots, I\}$. Then there exists a nonempty set $D(i,S)$ with the following three properties* (*see* (2.1.10), (2.5.3) *and* (2.11.1)):

$$D(i,S) \subset S; \tag{2.12.1}$$

$$\sum_{j \in D(i,S)} b_j \ge (\tau/2) \cdot \sum_{j \in S} b_j; \tag{2.12.2}$$

and

$$\max_{j,k \in D(i,S)} |h_{ij} - h_{ik}| \le \tau. \tag{2.12.3}$$

Proof. If $\sum_{j\in S} b_j = 0$, then this lemma is trivial (simply let $D(i,S) := \{j\}$ for some $j \in S$). Therefore, assume that $\sum_{j\in S} b_j > 0$.

Referring to (2.1.10), define the positive integer $K := 1/\tau$. For each $l \in \{1, 2, \ldots, 2K\}$, define the set

(2.12.4) $V(l) := \{j \in S : -1 - \zeta + (l-1)/K \le h_{ij} \le -1 - \zeta + l/K\}$.

By (2.11.7), one has that $S = \cup_{l=1}^{2K} V(l)$. Hence (see (2.5.6)) $\sum_{j\in S} b_j \le \sum_{l=1}^{2K} \sum_{j\in V(l)} b_j$. Hence there exists at least one element $l' \in \{1, 2, \ldots, 2K\}$ such that $\sum_{j\in V(l')} b_j \ge (2K)^{-1} \sum_{j\in S} b_j > 0$. Let $D(i,S) := V(l')$ for such an element l'. Then $D(i,S)$ is nonempty and satisfies (2.12.2) and, by (2.12.4), $D(i,S)$ also satisfies (2.12.1) and (2.12.3). This completes the proof of Lemma 2.12.

Definition 2.13. Let Δ denote the set of all integers $j \in \{1, \ldots, J\}$ with the following two properties (see (2.1.2) and (2.1.6)):

(2.13.1)
$$\sum_{i=1}^{I} \lambda_{ij}\varepsilon_{ij} \ge A_1 b_j;$$

and

(2.13.2)
$$\sum [r_{ij} \mid i \in \{1, \ldots, I\} : \lambda_{ij} = 0] \le A_5 b_j.$$

Lemma 2.14. *One has that* $\sum_{j\in\Delta} b_j \ge C/40$.

Proof. Define the sets

(2.14.1)
$$\Delta(1) := \Big\{ j \in \{1, \ldots, J\} : \sum_{i=1}^{I} \lambda_{ij}\varepsilon_{ij} \ge A_1 b_j \Big\}$$

and

(2.14.2)
$$\Delta(2) := \Big\{ j \in \{1, \ldots, J\} : \sum_{i=1}^{I} \lambda_{ij}\varepsilon_{ij} < A_1 b_j \Big\}.$$

Then by (2.5.6) and (2.1.2),

$$\sum_{j\in\Delta(2)}\sum_{i=1}^{I}\lambda_{ij}\varepsilon_{ij} \le \sum_{j\in\Delta(2)} A_1 b_j \le A_1 = C/10. \tag{2.14.3}$$

Hence by (2.11.8) and (2.14.1), $\sum_{j\in\Delta(1)}\sum_{i=1}^{I}\lambda_{ij}\varepsilon_{ij} \ge C/10$. Hence by (2.6.5) and (2.11.2), $\sum_{j\in\Delta(1)}\sum_{i=1}^{I} 2r_{ij} \ge C/10$. Hence by (2.6.4),

$$\sum_{j\in\Delta(1)} b_j \ge C/20. \tag{2.14.4}$$

Next define the set

(2.14.5)

$$\Delta(3) := \Big\{ j \in \{1,\dots,J\} : \sum [r_{ij} \,\big|\, i \in \{1,\dots,I\} : \lambda_{ij} = 0] > A_5 b_j \Big\}.$$

Then by (2.11.4) (and (2.6.2)),

$$A_6 \ge \sum_{j\in\Delta(3)} \sum [r_{ij} \,\big|\, i \in \{1,\dots,I\} : \lambda_{ij} = 0] \ge \sum_{j\in\Delta(3)} A_5 b_j.$$

Hence by (2.1.7),

$$\sum_{j\in\Delta(3)} b_j \le A_6/A_5 = C/40. \tag{2.14.6}$$

Now $\Delta(1) - \Delta(3) = \Delta$ by (2.14.1), (2.14.5) and Definition 2.13. Hence by (2.14.4) and (2.14.6) (and (2.5.6)),

$$\sum_{j\in\Delta} b_j \ge \sum_{j\in\Delta(1)} b_j - \sum_{j\in\Delta(3)} b_j \ge (C/20) - (C/40) = C/40.$$

Thus Lemma 2.14 holds.

Definition 2.15. Suppose S is a nonempty subset of the set Δ (from Definition 2.13). Let $U(S)$ denote the set of all $i \in \{1,\dots,I\}$ such that the following holds: There exists a pair of nonempty sets $T(i, S, +)$ and $T(i, S, -)$ with the following properties:

$$T(i, S, +) \subset S \quad \text{and} \quad T(i, S, -) \subset S; \tag{2.15.1}$$

$$\sum_{j \in T(i,S,+)} b_j \geq B_3 \sum_{j \in S} b_j; \tag{2.15.2}$$

$$\sum_{j \in T(i,S,-)} b_j \geq B_3 \sum_{j \in S} b_j; \tag{2.15.3}$$

$$\sup_{j,k \in T(i,S,+)} |h_{ij} - h_{ik}| \leq \tau; \tag{2.15.4}$$

$$\sup_{j,k \in T(i,S,-)} |h_{ij} - h_{ik}| \leq \tau; \tag{2.15.5}$$

$$\lambda_{ij} = 1 \quad \text{for all } j \in T(i, S, +); \tag{2.15.6}$$

and

$$\lambda_{ij} = -1 \quad \text{for all } j \in T(i, S, -). \tag{2.15.7}$$

Here, of course, B_3 is from (2.3.4). The next lemma involves constants from (2.3.5) and (2.1.4), as well as the numbers a_i and b_j from (2.5.2) and (2.5.3).

Lemma 2.16. *Suppose S is a subset of Δ such that*

$$\sum_{j \in S} b_j \geq B_4. \tag{2.16.1}$$

Then S is nonempty, and

$$\sum_{i \in U(S)} a_i \geq A_3. \tag{2.16.2}$$

Proof. As in the statement of the lemma, suppose $S \subset \Delta$ and (2.16.1) holds. Of course S is nonempty by (2.16.1) and (2.3.7).

The strategy for proving (2.16.2) will be as follows: First, some sets $Q(k)$, $1 \le k \le 7$, subsets of $\{1, \dots, I\}$, will be defined. Then it will be shown that $\sum_{i \in Q(7)} a_i \ge A_3$. (That will be most of the work in this proof.) Then at the end it will be shown that $Q(7) \subset U(S)$. Thereby (2.16.2) will be established.

Define the following seven subsets of $\{1, \dots, I\}$:

$$Q(1) := \Big\{ i \in \{1, \dots, I\} : \sum_{j \in S} \lambda_{ij} \varepsilon_{ij} \ge A_3 a_i \sum_{j \in S} b_j \Big\}; \tag{2.16.3}$$

$$Q(2) := \Big\{ i \in \{1, \dots, I\} : \sum_{j \in S} \lambda_{ij} \varepsilon_{ij} < A_3 a_i \sum_{j \in S} b_j \Big\}; \tag{2.16.4}$$

$$Q(3) := \Big\{ i \in Q(1) : \sum \big[|\varepsilon_{ij}| \,\big|\, j \in S : \lambda_{ij} = 0 \big] \ge A_4 a_i \sum_{j \in S} b_j \Big\}; \tag{2.16.5}$$

$$Q(4) := Q(1) - Q(3); \tag{2.16.6}$$

$$Q(5) := \Big\{ i \in Q(4) : \sum \big[b_j \,\big|\, j \in S : \lambda_{ij} = 1 \big] < B_2 \sum_{j \in S} b_j \Big\}; \tag{2.16.7}$$

$$Q(6) := \Big\{ i \in Q(4) : \sum \big[b_j \,\big|\, j \in S : \lambda_{ij} = -1 \big] < B_2 \sum_{j \in S} b_j \Big\}; \tag{2.16.8}$$

$$Q(7) := Q(4) - (Q(5) \cup Q(6)). \tag{2.16.9}$$

For convenience in the upcoming calculations, define the following numbers:

$$\mathcal{S} := \sum_{j \in S} b_j; \quad \text{and } \forall\, k \in \{1, \dots, 7\}, \quad \mathcal{Q}_k := \sum_{i \in Q(k)} a_i. \tag{2.16.10}$$

Note that by (2.16.1) and (2.3.7),

$$\mathcal{S} > 0. \tag{2.16.11}$$

Now by (2.11.2), (2.5.8), (2.1.1) and (2.16.10),

$$\sum_{i \in Q(1)} \sum_{j \in S} \lambda_{ij} \varepsilon_{ij} \le \sum_{i \in Q(1)} \sum_{j \in S} |\varepsilon_{ij}| \le \eta \cdot \mathcal{Q}_1 \cdot \mathcal{S}. \tag{2.16.12}$$

Also, by (2.16.4), (2.16.10), (2.5.6), (2.1.11) and (2.16.11),

$$\sum_{i\in Q(2)}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij}\le\sum_{i\in Q(2)}(A_3a_i\mathcal{S})\le A_3\mathcal{S}.$$

Adding this to (2.16.12) (see (2.16.3) and (2.16.4)), one has that

$$\sum_{i=1}^{I}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij}\le\eta\,\mathcal{Q}_1\,\mathcal{S}+A_3\mathcal{S}. \tag{2.16.13}$$

Also, by Definition 2.13 and the assumption (in the statement of Lemma 2.16) that $S\subset\Delta$, one has that $\sum_{i=1}^{I}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij}\ge A_1\mathcal{S}$. Combining this with (2.16.13) and dividing by $\mathcal{S}$ (see (2.16.11)), one has that $A_1\le\eta\,\mathcal{Q}_1+A_3$. Now $A_3\le A_1/3$ by (2.1.1), (2.1.3), (2.1.4) and (2.1.11). Hence by (2.1.3),

$$\mathcal{Q}_1\ge(A_1-A_3)/\eta\ge 2A_1/(3\eta)=2A_2. \tag{2.16.14}$$

Now let us look at the sets $Q(3)$ and $Q(4)$. By (2.16.10), (2.16.5), (2.6.5) and Definition 2.13, one has that

$$\begin{aligned}A_4\mathcal{Q}_3\mathcal{S}&=\sum_{i\in Q(3)}A_4a_i\mathcal{S}\le\sum_{i\in Q(3)}\sum[|\varepsilon_{ij}|\,\big|\,j\in S:\lambda_{ij}=0]\\&\le\sum_{i=1}^{I}\sum[|\varepsilon_{ij}|\,\big|\,j\in S:\lambda_{ij}=0]\\&\le\sum_{j\in S}\sum[2r_{ij}\,\big|\,i\in\{1,\dots,I\}:\lambda_{ij}=0]\\&\le\sum_{j\in S}2A_5b_j=2A_5\mathcal{S}.\end{aligned}$$

Hence by (2.1.6), (2.16.11) and (2.1.11), $\mathcal{Q}_3\le 2A_5/A_4=A_2$. Hence by (2.16.6) and (2.16.14),

$$\mathcal{Q}_4\ge\mathcal{Q}_1-\mathcal{Q}_3\ge A_2. \tag{2.16.15}$$

Next let us look at the set $Q(5)$. Now by (2.16.7) and (2.16.6), $Q(5)$ does not intersect $Q(3)$. Hence by (2.11.2), (2.5.8) and (2.16.7),

(2.16.16)
$$\begin{aligned}\sum_{i\in Q(5)}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij} &= -\sum_{i\in Q(5)}\sum_{j\in S}\varepsilon_{ij} + \sum_{i\in Q(5)}\sum[\varepsilon_{ij}\,|\,j\in S:\lambda_{ij}=0] \\ &\quad + 2\sum_{i\in Q(5)}\sum[\varepsilon_{ij}\,|\,j\in S:\lambda_{ij}=1] \\ &\le -\sum_{i\in Q(5)}\sum_{j\in S}\varepsilon_{ij} + \sum_{i\in Q(5)} A_4 a_i \mathcal{S} \\ &\quad + 2\sum_{i\in Q(5)}\sum[\eta a_i b_j\,|\,j\in S:\lambda_{ij}=1] \\ &\le -\sum_{i\in Q(5)}\sum_{j\in S}\varepsilon_{ij} + A_4\mathcal{Q}_5\mathcal{S} + 2\eta B_2\mathcal{Q}_5\mathcal{S}.\end{aligned}$$

Hence, by (2.5.9), (2.16.7)/(2.16.6)/(2.16.3), (2.3.3) and (2.16.1),

$$\begin{aligned}\delta_0 &\ge -\sum_{i\in Q(5)}\sum_{j\in S}\varepsilon_{ij} \ge \Big[\sum_{i\in Q(5)}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij}\Big] - A_4\mathcal{Q}_5\mathcal{S} - 2\eta B_2\mathcal{Q}_5\mathcal{S} \\ &\ge \Big[\sum_{i\in Q(5)} A_3 a_i\mathcal{S}\Big] - A_4\mathcal{Q}_5\mathcal{S} - 2\eta B_2\mathcal{Q}_5\mathcal{S} \\ &= (A_3 - A_4 - 2\eta B_2)\mathcal{Q}_5\mathcal{S} \\ &\ge (A_3 - A_4 - 2\eta B_2)\mathcal{Q}_5 B_4.\end{aligned}$$

Hence by (2.3.6), (2.3.3) and (2.3.7),

$$\mathcal{Q}_5 \le \frac{\delta_0}{(A_3 - A_4 - 2\eta B_2)\cdot B_4} \le B_1. \tag{2.16.17}$$

Next let us look at the set $Q(6)$. Now by (2.16.8) and (2.16.6), $Q(6)$

does not intersect $Q(3)$. Hence by (2.11.2), (2.5.8) and (2.16.8),

$$\begin{aligned}\sum_{i\in Q(6)}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij} &= \sum_{i\in Q(6)}\sum_{j\in S}\varepsilon_{ij} - \sum_{i\in Q(6)}\sum[\varepsilon_{ij}\,\big|\, j\in S:\lambda_{ij}=0]\\ &\quad - 2\sum_{i\in Q(6)}\sum[\varepsilon_{ij}\,\big|\, j\in S:\lambda_{ij}=-1]\\ &\le \sum_{i\in Q(6)}\sum_{j\in S}\varepsilon_{ij} + \sum_{i\in Q(6)} A_4 a_i\mathcal{S}\\ &\quad + 2\sum_{i\in Q(6)}\sum[a_i b_j\,\big|\, j\in S:\lambda_{ij}=-1]\\ &\le \sum_{i\in Q(6)}\sum_{j\in S}\varepsilon_{ij} + A_4\mathcal{Q}_6\mathcal{S} + 2B_2\mathcal{Q}_6\mathcal{S}.\end{aligned}$$

Hence by (2.5.9), (2.16.8)/(2.16.6)/(2.16.3), (2.3.3), (2.1.1) and (2.16.1),

$$\begin{aligned}\delta_0 &\ge \sum_{i\in Q(6)}\sum_{j\in S}\varepsilon_{ij} \ge \Big[\sum_{i\in Q(6)}\sum_{j\in S}\lambda_{ij}\varepsilon_{ij}\Big] - A_4\mathcal{Q}_6\mathcal{S} - 2B_2\mathcal{Q}_6\mathcal{S}\\ &\ge \Big[\sum_{i\in Q(6)} A_3 a_i\mathcal{S}\Big] - A_4\mathcal{Q}_6\mathcal{S} - 2B_2\mathcal{Q}_6\mathcal{S}\\ &= (A_3 - A_4 - 2B_2)\mathcal{Q}_6\mathcal{S}\\ &\ge (A_3 - A_4 - 2B_2)\mathcal{Q}_6 B_4.\end{aligned}$$

Hence by (2.3.6) and (2.3.3) and (2.3.7),

$$\mathcal{Q}_6 \le \frac{\delta_0}{(A_3 - A_4 - 2B_2)\cdot B_4} \le B_1. \tag{2.16.18}$$

Now by (2.16.9), (2.16.15), (2.16.17), (2.16.18) and (2.3.2),

$$\mathcal{Q}_7 \ge \mathcal{Q}_4 - (\mathcal{Q}_5 + \mathcal{Q}_6) \ge A_2 - 2B_1 \ge A_3. \tag{2.16.19}$$

Now refer to (2.16.2) (the conclusion of Lemma 2.16). By (2.5.6), (2.16.10) and (2.16.19), in order to prove (2.16.2) and thereby complete the proof of Lemma 2.16, it suffices to prove that

$$Q(7) \subset U(S). \tag{2.16.20}$$

Now in the rest of the proof of Lemma 2.16, we shall let $i \in \mathcal{Q}(7)$ be arbitrary but fixed. In order to show that $i \in U(S)$, and thereby prove (2.16.20) and complete the proof of Lemma 2.16, it suffices to produce two nonempty sets $T(i, S, +)$ and $T(i, S, -)$ such that equations (2.15.1)–(2.15.7) hold.

First note that since $i \in Q(7)$, one has by (2.16.9)/(2.16.8)/(2.16.7) that

$$\sum [b_j \,|\, j \in S : \lambda_{ij} = 1] \geq B_2 \mathcal{S}; \tag{2.16.21}$$

and

$$\sum [b_j \,|\, j \in S : \lambda_{ij} = -1] \geq B_2 \mathcal{S}. \tag{2.16.22}$$

Define the set $T(1) := \{j \in S : \lambda_{ij} = 1\}$. By (2.16.21), (2.3.7) and (2.16.11), $T(1)$ is nonempty. Applying Lemma 2.12, let $T(i, S, +) \subset T(1)$ be a nonempty set such that (2.15.4) holds and $\sum_{j \in T(i,S,+)} b_j \geq (\tau/2) \sum_{j \in T(1)} b_j$. Then $\sum_{j \in T(i,S,+)} b_j \geq (\tau/2) B_2 \mathcal{S} = B_3 \mathcal{S}$ by (2.16.21) and (2.3.4) (and (2.1.11)). Thus (2.15.2) holds. Also (2.15.6) and the first part of (2.15.1) hold by the definition of $T(1)$.

Define the set $T(2) := \{j \in S : \lambda_{ij} = -1\}$. By (2.16.22), (2.3.7) and (2.16.11), $T(2)$ is nonempty. Applying Lemma 2.12, let $T(i, S, -) \subset T(2)$ be a nonempty set such that (2.15.5) holds and $\sum_{j \in T(i,S,-)} b_j \geq (\tau/2) \sum_{j \in T(2)} b_j$. Then $\sum_{j \in T(i,S,-)} b_j \geq (\tau/2) B_2 \mathcal{S} = B_3 \mathcal{S}$ by (2.16.22) and (2.3.4). Thus (2.15.3) holds. Also (2.15.7) and the second part of (2.15.1) hold by the definition of $T(2)$.

The sets $T(i, S, +)$ and $T(i, S, -)$ are nonempty and satisfy equations (2.15.1)–(2.15.7). Thus, $i \in U(S)$ (by Definition 2.15). Since i was an arbitrary element of $Q(7)$, equation (2.16.20) holds. This completes the proof of Lemma 2.16.

Definition 2.17. Refer to (2.11.1). For every integer $i \in \{1, \ldots, I\}$ and every nonempty set $S \subset \{1, \ldots, J\}$, define the number

$$H(i, S) := \min_{j \in S} h_{ij}. \tag{2.17.1}$$

(The use of "min" is only for definiteness. For what follows, it is important only that $H(i, S)$ be one of the numbers h_{ij}, $j \in S$.)

Step 2.18. Recall from Step 2.5 the notations p_{ij}, a_i and b_j related to the probability space $(\Omega, \mathcal{M}, P)$ and the σ-fields $\mathcal{F}$ and $\mathcal{G} \subset \mathcal{M}$ from Step 2.4. We will need corresponding notations related to the "other" probability space $(\Omega^*, \mathcal{M}^*, P^*)$ in Steps 2.2 and 2.3.

Refer to Step 2.2 and the first paragraph of Step 2.3. For each $m \in \{1, \dots, M\}$ and each $n \in \{1, \dots, N\}$, define the nonnegative number

$$p^*_{mn} := P^*(F^*_m \cap G^*_n). \tag{2.18.1}$$

For each $m \in \{1, \dots, M\}$, define the nonnegative number

$$a^*_m := P^*(F^*_m) = \sum_{n=1}^{N} p^*_{mn}. \tag{2.18.2}$$

For each $n \in \{1, \dots, N\}$, define the nonnegative number

$$b^*_n := P^*(G^*_n) = \sum_{m=1}^{M} p^*_{mn}. \tag{2.18.3}$$

Step 2.19. Refer again to the positive integers M and N from Step 2.3 and the positive integer I from Step 2.4.

For a given sequence, $z_1, z_2, \dots, z_M$ of elements of $\{1, \dots, I\}$, we would like to recursively define a two-dimensional array of sets indexed by $m \in \{1, \dots, M\}$ and $n \in \{1, \dots, N\}$. The "syntax" of the recursive definition will be slightly less cumbersome if we allow the use of a superfluous "initial" coordinate $z_0 \in \{1, \dots, I\}$.

In the recursive definition (given below) we shall define for each $m \in \{0, 1, \dots, M\}$ and each vector $\mathbf{y} := (z_0, z_1, \dots, z_m) \in \{1, \dots, I\}^{m+1}$ a collection of sets $S(m, n, \mathbf{y}) \subset \Delta$, $n \in \{1, \dots, N\}$ such that for each $n \in \{1, \dots, N\}$,

$$\sum_{j \in S(m,n,\mathbf{y})} b_j \geq B_3^m C/40. \tag{2.19.1}$$

(Here of course the set Δ is from Definition 2.13.) Also, for each $m \in \{1, \dots, M\}$ (but not $m = 0$), and each $\mathbf{y} := (z_0, z_1, \dots, z_m) \in$

$\{1, \ldots, I\}^{m+1}$, we shall define a pair of sets $\mathcal{D}(m, \mathbf{y})$ and $\mathcal{E}(m, \mathbf{y}) \subset \{1, \ldots, N\}$ and also a number $\sigma(m, \mathbf{y}) \in \{-1, 1\}$.

The use of the vector $\mathbf{y} := (z_0, z_1, \ldots, z_m) \in \{1, \ldots, I\}^{m+1}$ (instead of a vector $(z_0, z_1, \ldots, z_M) \in \{1, \ldots, I\}^{M+1}$) may seem awkward but will facilitate some arguments later on.

Let us start with $m = 0$. For each $\mathbf{y} := (z_0) \in \{1, \ldots, I\}$ and each $n \in \{1, \ldots, N\}$, define the set $S(0, n, \mathbf{y}) := \Delta$ (from Definition 2.13). By Lemma 2.14 (and (2.3.7)), equation (2.19.1) holds with $m = 0$.

Now suppose $m \in \{1, \ldots, M\}$, and suppose the set $S(m-1, n, \mathbf{w}) \subset \Delta$ has already been defined for each $\mathbf{w} = (z_0, \ldots, z_{m-1}) \in \{1, \ldots, I\}^m$ and each $n \in \{1, \ldots, N\}$, and that (for each such $\mathbf{w}$ and n), (2.19.1) holds with m replaced by $m-1$ and $\mathbf{y}$ replaced by $\mathbf{w}$:

$$\sum_{j \in S(m-1, n, \mathbf{w})} b_j \geq B_3^{m-1} C/40. \tag{2.19.2}$$

Note that by (2.19.2), (2.3.7) and (2.1.1), the set $S(m-1, n, \mathbf{w})$ is nonempty for each $\mathbf{w}$ and each n.

Suppose $\mathbf{y} := (y_0, y_1, \ldots, y_m) \in \{1, \ldots, I\}^{m+1}$. Denote $\mathbf{x} := (y_0, y_1, \ldots, y_{m-1})$, the vector $\in \{1, \ldots, I\}^m$ consisting of the first m coordinates of $\mathbf{y}$.

Referring to Definition 2.15, define the sets

$$\mathcal{D}(m, \mathbf{y}) := \big\{ n \in \{1, \ldots, N\} : y_m \notin U(S(m-1, n, \mathbf{x})) \big\}; \tag{2.19.3}$$

and

$$\mathcal{E}(m, \mathbf{y}) := \big\{ n \in \{1, \ldots, N\} : y_m \in U(S(m-1, n, \mathbf{x})) \big\}. \tag{2.19.4}$$

Either one of these two sets could be empty. Obviously, these two sets complement each other in $\{1, \ldots, N\}$.

For each $n \in \mathcal{D}(m, \mathbf{y})$ (if the set $\mathcal{D}(m, \mathbf{y})$ is nonempty), define the set

$$S(m, n, \mathbf{y}) := D(y_m, S(m-1, n, \mathbf{x})) \tag{2.19.5}$$

from Lemma 2.12. For each $n \in \mathcal{D}(m, \mathbf{y})$, one has that (i) $S(m, n, \mathbf{y}) \subset \Delta$, and (ii) (2.19.1) holds. To see (i), note that $S(m, n, \mathbf{y}) \subset S(m-1, n, \mathbf{x}) \subset \Delta$ (see (2.12.1)). To see (ii), first note that $B_3 < \tau/2$ by (2.3.4) and (2.3.7), and then apply (2.12.2) and (2.19.2). Note that for each $n \in \mathcal{D}(m, \mathbf{y})$ (if $\mathcal{D}(m, \mathbf{y})$ is nonempty), the set $S(m, n, \mathbf{y})$ is nonempty, by (2.19.1), (2.1.1) and (2.3.7).

Next, referring to (2.17.1) and (2.18.1)/(2.18.2)/(2.18.3), define the number $\sigma(m, \mathbf{y}) \in \{-1, 1\}$ as follows:

$$\sigma(m, \mathbf{y}) := \begin{cases} 1 & \text{if } \sum_{n \in \mathcal{D}(m,\mathbf{y})} H(y_m, S(m, n, \mathbf{y})) \\ & \quad \cdot (p^*_{mn} - a^*_m b^*_n) \geq 0 \\ -1 & \text{if } \sum_{n \in \mathcal{D}(m,\mathbf{y})} H(y_m, S(m, n, \mathbf{y})) \\ & \quad \cdot (p^*_{mn} - a^*_m b^*_n) < 0. \end{cases} \tag{2.19.6}$$

(Of course, if $\mathcal{D}(m, \mathbf{y})$ is empty, then the sum is 0 and $\sigma(m, \mathbf{y}) = 1$.)

Next refer to (2.19.4) and Definition 2.15. For each $n \in \mathcal{E}(m, \mathbf{y})$ (if the set $\mathcal{E}(m, \mathbf{y})$ is nonempty), define the following set from Definition 2.15:

(2.19.7)

$$S(m, n, \mathbf{y}) := \begin{cases} T(y_m, S(m-1, n, \mathbf{x}), +) & \text{if } \sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) \geq 0 \\ T(y_m, S(m-1, n, \mathbf{x}), -) & \text{if } \sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) < 0. \end{cases}$$

For each $n \in \mathcal{E}(m, \mathbf{y})$, one has that (i) $S(m, n, \mathbf{y}) \subset \Delta$, and (ii) equation(2.19.1) holds. Here (i) holds by (2.15.1) since $S(m-1, n, \mathbf{x}) \subset \Delta$; and (ii) holds by (2.19.2) and (2.15.2)/(2.15.3).

This completes the recursive definition (for $m = 1, \dots, M$) of the sets $S(m, n, \mathbf{y})$ ($\subset \Delta$ and satisfying (2.19.1)) and the sets $\mathcal{D}(m, \mathbf{y})$ and $\mathcal{E}(m, \mathbf{y})$, and the number $\sigma(m, \mathbf{y})$, for $\mathbf{y} \in \{1, \dots, I\}^{m+1}$ and $n \in \{1, \dots, N\}$.

Lemma 2.20. *Suppose $m \in \{1, \dots, M\}$, $\mathbf{y} := (y_0, y_1, \dots, y_m) \in \{1, \dots, I\}^{m+1}$ and $n \in \{1, \dots, N\}$. Then the following four statements hold*:

(1) *One has that $\sum_{j \in S(m,n,\mathbf{y})} b_j \geq B_4$.*

(2) *The set $S(m, n, \mathbf{y})$ is nonempty and, for every $j \in S(m, n, \mathbf{y})$, one has that* (*referring to Definition* 2.17 *and writing y_m also as $y(m)$ and h_{ij} also as $h_{i,j}$*)

$$|h_{y(m),j} - H(y_m, S(m, n, \mathbf{y}))| \leq \tau. \tag{2.20.1}$$

(3) *Letting* $\mathbf{x} := (y_0, y_1, \dots, y_{m-1}) \in \{1, \dots, I\}^m$ *(that is,* $\mathbf{x}$ *consists of the first* m *coordinates of* $\mathbf{y}$*), one has that*

$$S(m,n,\mathbf{y}) \subset S(m-1,n,\mathbf{x}) \subset \Delta. \tag{2.20.2}$$

(4) *If* $n \in \mathcal{E}(m,\mathbf{y})$*, then*

$$H(y_m, S(m,n,\mathbf{y})) \cdot \sigma(m,\mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) \geq A_7 \cdot |p^*_{mn} - a^*_m b^*_n|. \tag{2.20.3}$$

Proof. Statement (1) holds by (2.19.1), (2.3.5), (2.1.1) and (2.3.7).

In the rest of this proof, in the manner of statement (2), y_m is also written as $y(m)$, and h_{ij} also as $h_{i,j}$ and λ_{ij} also as $\lambda_{i,j}$.

Proof of (2). The set $S(m,n,\mathbf{y})$ is nonempty by (say) statement (1) and (2.3.7). Also,

$$\max_{j,k \in \mathcal{S}(m,n,\mathbf{y})} |h_{y(m),j} - h_{y(m),k}| \leq \tau. \tag{2.20.4}$$

If $n \in \mathcal{D}(m,\mathbf{y})$, then (2.20.4) holds by (2.19.5) and (2.12.3). If instead $n \in \mathcal{E}(m,\mathbf{y})$, then (2.20.4) holds by (2.19.7) and (2.15.4)/(2.15.5).

Now from (2.17.1), $H(y_m, S(m,n,\mathbf{y}))$ is one of the numbers $h_{y(m),k}$, $k \in S(m,n,\mathbf{y})$. Hence (2.20.1) follows from (2.20.4). This completes the proof of (2).

Proof of (3). This was built into the recursive definition in Step 2.19. (If $n \in \mathcal{D}(m,\mathbf{y})$, then the first part of (2.20.2) holds by (2.19.5) and (2.12.1). If instead $n \in \mathcal{E}(m,\mathbf{y})$, then the first part of (2.20.2) holds by (2.19.7) and (2.15.1). The "$\subset \Delta$" in (2.20.2) holds from Step 2.19 regardless of whether $m-1=0$ or $m-1 \geq 1$.)

Proof of (4). Suppose $n \in \mathcal{E}(m,\mathbf{y})$. Let $\mathbf{x} := (y_0, y_1, \dots, y_{m-1})$ as in statement (3).

Note that $|\sigma(m,\mathbf{y})| = 1$ by (2.19.6).

The argument will be broken into two cases:

Case I. $\sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) \geq 0$. Then

$$\begin{aligned} \sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) &= |\sigma^*(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n)| \\ &= |p^*_{mn} - a^*_m b^*_n|. \end{aligned} \tag{2.20.5}$$

Now referring to (2.17.1), let $j \in S(m, n, \mathbf{y})$ be such that

$$H(y_m, S(m, n, \mathbf{y})) = h_{y(m),j}. \tag{2.20.6}$$

Then $j \in T(y_m, S(m-1, n, \mathbf{x}), +)$ by (2.19.7); hence $\lambda_{y(m),j} = 1$ by (2.15.6), hence $h_{y(m),j} \geq A_7$ by (2.11.2). Hence by (2.20.5) and (2.20.6), equation (2.20.3) holds.

Case II. $\sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) < 0$. Then

$$\begin{aligned} -\sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n) &= |\sigma(m, \mathbf{y}) \cdot (p^*_{mn} - a^*_m b^*_n)| \\ &= |p^*_{mn} - a^*_m b^*_n|. \end{aligned} \tag{2.20.7}$$

Again referring to (2.17.1), let $j \in S(m, n, \mathbf{y})$ be such that (2.20.6) holds. Then $j \in T(y_m, S(m-1, n, \mathbf{x}), -)$ by (2.19.7), $\lambda_{y(m),j} = -1$ by (2.15.7), $-h_{y(m),j} \geq A_7$ by (2.11.2), and hence (2.20.3) holds by (2.20.6) and (2.20.7). This completes the proof of statement (4) and of Lemma 2.20.

Step 2.21. In connection with equations (2.19.3) and (2.19.4), the following "indicator functions" will be useful later on:

For each $m \in \{1, \ldots, M\}$, each $\mathbf{y} := (y_0, y_1, \ldots, y_m) \in \{1, \ldots, I\}^{m+1}$ and each $n \in \{1, \ldots, N\}$, define

$$\mathcal{I}(m, n, \mathbf{y}) := \begin{cases} 1 & \text{if } n \in \mathcal{E}(m, \mathbf{y}), \\ 0 & \text{if } n \in \mathcal{D}(m, \mathbf{y}). \end{cases} \tag{2.21.1}$$

Step 2.22. For any $m \in \{1, \ldots, M\}$, any vector $\mathbf{y} := (y_0, y_1, \ldots, y_m) \in \{1, \ldots, I\}^{m+1}$, and any $n \in \{1, \ldots, N\}$, define (for convenient notation) the number

$$W(m, n, \mathbf{y}) := \sigma(m, \mathbf{y}) \cdot H(y_m, S(m, n, \mathbf{y})) \cdot [p^*_{mn} - a^*_m b^*_n]. \tag{2.22.1}$$

(This is the left side of (2.20.3).)

For any given vector $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \{1, \dots, I\}^{M+1}$, we shall use, for each $m \in \{0, 1, \dots, M\}$, the notation $\mathbf{z}_m := (z_0, z_1, \dots, z_m) \in \{1, \dots, I\}^{m+1}$. (Thus $\mathbf{z}_M = \mathbf{z}$.) For a given $m \in \{0, 1, \dots, M\}$, the vector $\mathbf{z}_m$ will sometimes be written $\mathbf{z}(m)$.

For each $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \{1, \dots, I\}^{M+1}$, one has by (2.19.6) that

$$\sum_{m=1}^{M} \sum_{n \in \mathcal{D}(m,\mathbf{z}(m))} W(m, n, \mathbf{z}_m) \geq 0;$$

and also by Lemma 2.20(4),

$$\sum_{m=1}^{M} \sum_{n \in \mathcal{E}(m,\mathbf{z}(m))} W(m, n, \mathbf{z}_m) \geq \sum_{m=1}^{M} \sum_{n \in \mathcal{E}(m,\mathbf{z}(m))} A_7 \cdot |p^*_{mn} - a^*_m b^*_n|.$$

Hence by (2.19.3)/(2.19.4) and (2.21.1), for each $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \{1, \dots, I\}^{M+1}$,

$$\sum_{m=1}^{M} \sum_{n=1}^{N} W(m, n, \mathbf{z}_m) \geq A_7 \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} \mathcal{I}(m, n, \mathbf{z}_m) \cdot |p^*_{mn} - a^*_m b^*_n|. \tag{2.22.2}$$

Lemma 2.23. *There exists a vector* $\mathbf{z}^* := (z^*_0, z^*_1, \dots, z^*_M) \in \{1, \dots, I\}^{M+1}$ *such that*

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \mathcal{I}(m, n, \mathbf{z}^*_m) \cdot |p^*_{mn} - a^*_m b^*_n| \geq A_3 \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} |p^*_{mn} - a^*_m b^*_n|. \tag{2.23.1}$$

Proof. The easiest way to prove this lemma seems to be an elementary "probabilistic" argument. In carrying out this argument, we need to avoid conflicting with the notations for the probability spaces defined in Steps 2.2 and 2.4.

Define the probability space $(\Omega^{**}, \mathcal{M}^{**}, P^{**})$ as follows: $\Omega^{**} := \{1, \dots, I\}^{M+1}$, $\mathcal{M}^{**}$ is the σ-field consisting of all subsets of Ω^{**}, and P^{**} is the measure on $(\Omega^{**}, \mathcal{M}^{**})$ such that for each $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \Omega^{**}$,

$$P^{**}(\{\mathbf{z}\}) = \prod_{m=0}^{M} a(z_m), \tag{2.23.2}$$

where $a(i) := a_i$ from (2.5.2). By (2.5.6) and a trivial argument, P^{**} is a probability measure on $(\Omega^{**}, \mathcal{M}^{**})$.

Define the random variables $Z_0, Z_1, \dots, Z_m$ on Ω^{**} as follows. For each $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \Omega^{**}$ and each $m \in \{0, 1, \dots, M\}$,

$$Z_m(\mathbf{z}) := z_m. \tag{2.23.3}$$

For each $m \in \{0, 1, \dots, M\}$, this random variable takes its values in $\{1, \dots, I\}$. By (2.23.2), (2.23.3), (2.5.6) and a trivial calculation, for each $m \in \{0, 1, \dots, M\}$ and each $i \in \{1, \dots, I\}$,

$$P^{**}(Z_m = i) = a_i. \tag{2.23.4}$$

As a simple consequence of (2.23.2), (2.23.3) and (2.23.4), the random variables $Z_0, Z_1, \dots, Z_M$ are independent.

Define the random vector $\mathbf{Z} := (Z_0, Z_1, \dots, Z_M)$. For each $m \in \{0, 1, \dots, m\}$, define the random vector $\mathbf{Z}_m := (Z_0, Z_1, \dots, Z_m)$. (Then $\mathbf{Z}_M = \mathbf{Z}$.) The following observation will be handy:

$$\begin{aligned}&\text{For each } m \in \{1, \dots, M\}, \text{ the random variable}\\ &Z_m \text{ is independent of the random vector } \mathbf{Z}_{m-1}.\end{aligned} \tag{2.23.5}$$

The following observation will also be useful:

$$\begin{aligned}&\text{If } S \subset \Delta \text{ and } \sum_{j \in S} b_j \geq B_4, \text{ then } \forall\, m \in \{0, 1, \dots, M\},\\ &P^{**}(Z_m \in U(S)) \geq A_3.\end{aligned} \tag{2.23.6}$$

This holds because, by (2.23.4) and Lemma 2.16,

$$P^{**}(Z_m \in U(S)) = \sum_{i \in U(S)} P^{**}(Z_m = i) = \sum_{i \in U(S)} a_i \geq A_3.$$

In what follows, the "expected value" of a random variable X on Ω^{**} will be denoted $E^{**}X$.

Refer to Step 2.19. For each vector $\mathbf{z} := (z_0, z_1, \dots, z_M) \in \{1, \dots, I\}^{M+1}$, each $m \in \{0, 1, \dots, M\}$ and each $n \in \{1, \dots, N\}$, one has that $S(m, n, \mathbf{z}_m) \subset \Delta$ and $\sum_{j \in S(m,n,\mathbf{z}(m))} b_j \geq B_4$, by Lemma 2.20 (1), (3) (or for $m = 0$, by (2.3.5), (2.3.7), (2.19.1) and the definition $S(0, n, \mathbf{y}) = \Delta$ in Step 19). With that in mind, in the calculations just below, *each sum is taken over all sets* $S \subset \Delta$ *such that* $\sum_{j \in S} b_j \geq B_4$.

For each $m \in \{1, \dots, M\}$ and each $n \in \{1, \dots, N\}$, by (2.21.1), (2.19.4), (2.23.5) and (2.23.6),

$$\begin{aligned}
E^{**}\mathcal{I}(m, n, \mathbf{Z}_m) &= P^{**}(n \in \mathcal{E}(m, \mathbf{Z}_m)) \\
&= P^{**}\big(Z_m \in U(S(m-1, n, \mathbf{Z}_{m-1}))\big) \\
&= \sum_S P^{**}\big(Z_m \in U(S(m-1, n, \mathbf{Z}_{m-1})) \text{ and} \\
&\qquad\qquad S(m-1, n, \mathbf{Z}_{m-1}) = S\big) \\
&= \sum_S P^{**}\big(Z_m \in U(S) \text{ and } S(m-1, n, \mathbf{Z}_{m-1}) = S\big) \\
&= \sum_S P^{**}(Z_m \in U(S)) \cdot P^{**}(S(m-1, n, \mathbf{Z}_{m-1}) = S) \\
&\geq \sum_S A_3 \cdot P^{**}(S(m-1, n, \mathbf{Z}_{m-1}) = S) \\
&= A_3 \cdot 1.
\end{aligned}$$

Hence,

$$\begin{aligned}
&E^{**} \sum_{m=1}^{M} \sum_{n=1}^{N} \mathcal{I}(m, n, \mathbf{Z}_m) \cdot |p^*_{mn} - a^*_m b^*_n| \\
&\qquad = \sum_{m=1}^{M} \sum_{n=1}^{N} |p^*_{mn} - a^*_m b^*_n| \cdot E^{**}\mathcal{I}(m, n, \mathbf{Z}_m) \\
&\qquad \geq \sum_{m=1}^{M} \sum_{n=1}^{N} |p^*_{mn} - a^*_m b^*_n| \cdot A_3.
\end{aligned} \tag{2.23.7}$$

Now for (say) a random variable defined on a probability space with only finitely many elements (as in Ω^{**}), the expected value cannot be greater than the maximum value taken by that random variable. Hence by (2.23.7), there has to exist an element $\mathbf{z}^* := (z_0^*, z_1^*, \dots, z_M^*) \in \Omega^{**}$ such that

$$\sum_{m=1}^{M}\sum_{n=1}^{N} \mathcal{I}(m, n, \mathbf{Z}_m(\mathbf{z}^*)) \cdot |p_{mn}^* - a_m^* b_n^*| \geq A_3 \cdot \sum_{m=1}^{M}\sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|.$$

Since $\mathbf{Z}_m(\mathbf{z}^*) = (z_0^*, z_1^*, \dots, z_m^*) = \mathbf{z}_m^*$ for each $m = 1, \dots, M$ by (2.23.3), one has that equation (2.23.1) holds. This completes the proof of Lemma 2.23.

Step 2.24. In this final step, the numbers h_{ij} in (2.11.1) will also be written as $h(i, j)$, and the elements $\theta_i \in \Theta$ and $\gamma_j \in \Gamma$ from (2.4.2) will also be written as $\theta(i)$ and $\gamma(j)$.

Applying Lemma 2.23, henceforth let $\mathbf{z}^* := (z_0^*, z_1^*, \dots, z_M^*) \in \{1, \dots, I\}^{M+1}$ be fixed such that (2.23.1) holds. As usual, for each $m \in \{0, 1, \dots, M\}$, we let $\mathbf{z}_m^* := (z_0^*, z_1^*, \dots, z_m^*)$. (Then $\mathbf{z}_M^* = \mathbf{z}$.)

By (2.22.2), (2.23.1) and (2.1.9), one has that

$$\sum_{m=1}^{M}\sum_{n=1}^{N} W(m, n, \mathbf{z}_m^*) \geq A_8 \cdot \sum_{m=1}^{M}\sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|. \tag{2.24.1}$$

For each $n \in \{1, \dots, N\}$, let j_n be an element of the (nonempty) set $S(M, n, \mathbf{z}^*)$ (see Lemma 2.20 (2)). Then by Lemma 2.20 (3) for each $n \in \{1, \dots, N\}$,

$$\begin{aligned} j_n \in S(M, n, \mathbf{z}^*) &\subset S(M-1, n, \mathbf{z}_{M-1}^*) \\ &\subset \cdots \subset S(1, n, \mathbf{z}_1^*) \subset S(0, n, \mathbf{z}_0^*) = \Delta. \end{aligned} \tag{2.24.2}$$

(Here the last equality comes from Step 2.19.) Hence by Lemma 2.20 (2), for each $m \in \{1, \dots, M\}$ and each $n \in \{1, \dots, N\}$,

$$|h(z_m^*, j_n) - H(z_m^*, S(m, n, \mathbf{z}_m^*))| \leq \tau. \tag{2.24.3}$$

Since $|\sigma(m, \mathbf{z}_m^*)| = 1$ for each $m \in \{1, \dots, M\}$ (see (2.19.6)), one has that

$$\begin{aligned} &\sum_{m=1}^{M}\sum_{n=1}^{N} |\sigma(m, \mathbf{z}_m^*)| \cdot |h(z_m^*, j_n) - H(z_m^*, S(m, n, \mathbf{z}_m^*))| \cdot |p_{mn}^* - a_m^* b_n^*| \\ &\leq \tau \cdot \sum_{m=1}^{M}\sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|. \end{aligned} \tag{2.24.4}$$

Since $A_8 - \tau \geq \tau$ by (2.1.10), one has by (2.22.1), (2.24.1) and (2.24.4) that

$$\begin{aligned} &\sum_{m=1}^{M}\sum_{n=1}^{N} \sigma(m, \mathbf{z}_m^*) \cdot h(z_m^*, j_n) \cdot [p_{mn}^* - a_m^* b_n^*] \\ &\qquad \geq \tau \cdot \sum_{m=1}^{M}\sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|. \end{aligned} \tag{2.24.5}$$

Next, for each $m \in \{1, \dots, M\}$, by (2.18.1), (2.18.2), (2.18.3) and trivial calculations,

$$\sum_{n=1}^{N}[p_{mn}^* - a_m^* b_n^*] = \Big(\sum_{n=1}^{N} p_{mn}^*\Big) - \Big(a_m^* \sum_{n=1}^{N} b_n^*\Big) = a_m^* - a_m^* \cdot 1 = 0.$$

Hence, referring to Step 2.11, one has by (2.24.5) that

$$\begin{aligned} &\sum_{m=1}^{M}\sum_{n=1}^{N} \sigma(m, \mathbf{z}_m^*) \cdot [h(z_m^*, j_n) + \zeta] \cdot [p_{mn}^* - a_m^* b_n^*] \\ &\qquad \geq \tau \cdot \sum_{m=1}^{M}\sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|. \end{aligned} \tag{2.24.6}$$

By (2.11.1) and (2.6.6), for each $m \in \{1, \dots, M\}$ and each $n \in \{1, \dots, N\}$,

$$h(z_m^*, j_n) + \zeta = \chi(\theta(z_m^*), \gamma(j_n)). \tag{2.24.7}$$

For each $n \in \{1, \dots, N\}$, define the element $\gamma_n^* \in \Gamma$ by $\gamma_n^* := \gamma(j_n)$. For each $m \in \{1, \dots, M\}$, define the element $\theta_m^* \in \Theta$ as follows:

$$\theta_m^* := \begin{cases} \theta(z_m^*) & \text{if } \sigma(m, \mathbf{z}_m^*) = 1 \\ \tilde{\theta}(z_m^*) & \text{if } \sigma(m, \mathbf{z}_m^*) = -1. \end{cases} \tag{2.24.8}$$

Here $\tilde{\theta}(z_m^*)$ is an element of Θ such that (see (2.2.1)), $\chi(\tilde{\theta}(z_m^*), \gamma) = -\chi(\theta(z_m^*), \gamma)$ for all $\gamma \in \Gamma$.

For each $m \in \{1, \ldots, M\}$ and each $n \in \{1, \ldots, N\}$ one has by (2.24.8) and (2.24.7) that

$$\begin{aligned}\sigma(m, \mathbf{z}_m^*) \cdot [h(z_m^*, j_n) + \zeta] &= \sigma(m, \mathbf{z}_m^*) \cdot \chi(\theta(z_m^*), \gamma(j_n)) \\ &= \chi(\theta_m^*, \gamma_n^*),\end{aligned} \tag{2.24.9}$$

regardless of whether $\sigma(m, \mathbf{z}_m^*) = 1$ or $\sigma(m, \mathbf{z}_m^*) = -1$. Substituting this into (2.24.6), one has that

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \chi(\theta_m^*, \gamma_n^*) \cdot [p_{mn}^* - a_m^* b_n^*] \geq \tau \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*|.$$

Hence by (1.11) (in Section 1), the first paragraph of Step 2.3, equations (2.18.1)/(2.18.2)/(2.18.3) and (2.3.1),

$$R_{\Theta,\Gamma,\chi}(\mathcal{F}^*, \mathcal{G}^*) \geq \tau \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} |p_{mn}^* - a_m^* b_n^*| \geq \tau \cdot \beta(\mathcal{F}^*, \mathcal{G}^*).$$

Hence (2.2.2) holds. This completes the proof of Theorem 1.8.

Acknowledgment. This work was inspired by conversations years ago with Włodzimierz Bryc and Svante Janson. The author also thanks Grahame Bennett for helpful comments.

REFERENCES

1. C. Bennett and R. Sharpley, *Interpolation of operators*, Academic Press, Boston, 1988.

2. J. Bergh and J. Löfström, *Interpolation spaces*, Springer-Verlag, New York, 1976.

3. R.C. Bradley, *On tightness of partial sums from strictly stationary, absolutely regular sequences of B-valued random variables*, J. Math. Anal. Appl. **240** (1999), 128–144.

4. R.C. Bradley and W. Bryc, *Multilinear forms and measures of dependence between random variables*, J. Multivariate Anal. **16** (1985), 335–367.

5. R.C. Bradley, W. Bryc and S. Janson, *On dominations between measures of dependence*, J. Multivariate Anal. **23** (1987), 312–329.

6. ———, *Remarks on the foundations of measures of dependence*, in *New perspectives in theoretical and applied statistics* (M.L. Puri, J.P. Vilaplana and W. Wertz, eds.), Wiley, New York, 1987, pp. 421–437.

7. H. Dehling and W. Philipp, *Almost sure invariance principles for weakly dependent vector-valued random variables*, Ann. Probab. **10** (1982), 689–701.

8. P. Doukhan, P. Massart and E. Rio, *The functional central limit theorem for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), 63–82.

9. I.A. Ibragimov and Yu.V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff, Groningen, 1971.

10. F. Merlevède, M. Peligrad and S.A. Utev, *Sharp conditions for the CLT of linear processes in a Hilbert space*, J. Theoret. Probab. **10** (1997), 681–693.

11. W. Philipp, *Weak and L^p-invariance principles for sums of B-valued random variables*, Ann. Probab. **8** (1980), 68–82, (correction ibid. **14** (1986), 1095–1101).

12. M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U.S.A. **42** (1956), 43–47.

13. ———, *Markov processes, structure and asymptotic behavior*, Springer-Verlag, Berlin, 1971.

14. P. Shields, *The theory of Bernoulli shifts*, University of Chicago Press, Chicago, 1973.

15. V.A. Volkonskii and Yu.A. Rozanov, *Some limit theorems for random functions*, I, Theory Probab. Appl. **4** (1959), 178–197.

Department of Mathematics, Indiana University, Bloomington, Indiana 47405
E-mail address: bradleyr@indiana.edu