THE INCLUSION PROBLEM FOR MIXED-NORM SPACES

WAYNE GREY AND GORD SINNAMON

Abstract. Given two mixed norm Lebesgue spaces on an n-fold product of arbitrary σ-finite measure spaces, when is one contained in the other? If so, what is the norm of the inclusion map? These questions are answered completely for a large range of Lebesgue indices and all measure spaces. When the measure spaces are atomless both questions are settled for all indices. When the measure spaces are not purely atomic the first question is settled for all indices. Some complete and some partial results are given in the remaining cases, but a wide variety of behaviour is observed. In particular, the norm problem for purely atomic measure spaces is seen to be intractable for certain ranges of the Lebesgue indices; it is equivalent to an optimization problem that includes a known NP-hard problem as a special case.

1. Introduction

Mixed norm spaces are vector spaces of multivariable functions equipped with norms that take advantage of the product structure on the domain by successively applying different norms with respect to each variable. Mixed norm Lebesgue spaces, where the norms on each factor are L^p norms, were introduced and studied by Benedek and Panzone in [3]; they defined,

$$\|f\|_P = \left(\int \dots \left(\int \left(\int |f(x_1, x_2, \dots, x_n)|^{p_1}\, d\mu_1\right)^{p_2/p_1} d\mu_2\right)^{p_3/p_2} \dots\, d\mu_n\right)^{1/p_n}$$

where $P = (p_1, p_2, \dots, p_n)$. Mixed norm spaces of this type arise naturally in harmonic and functional analysis. The particular special case of amalgam spaces are used to define and study function spaces in which the local and global behavior are independent of one another. Mixed norm spaces based on Lorentz spaces instead of Lebesgue spaces have been used to study Fourier and Sobolev inequalities in [1, 2, 11, 12, 15], see [8] for recent work and additional references. Mixed norms in which the norm on each factor is a general Banach function norm were studied in [5] and [6]. For applications to p-summing operators and the Bohnenblust-Hille inequality, see [9, 10, 16, 17].

In this article we consider the fundamental inclusion problem: *Given two mixed norm Lebesgue spaces when is one a subset of the other?*

In the single-variable case the solution to the inclusion problem is well known. But the solution depends crucially on the atomic structure of the underlying measures. For mixed norms we will see that for atomless measures the inclusion problem

2010 *Mathematics Subject Classification.* Primary 46E30, Secondary 46A45, 26D15.

Key words and phrases. mixed-norm space, embedding, Minkowski inequality, product measure, atomic measure, partition problem.

Supported by the Natural Sciences and Engineering Research Council of Canada.

can be answered completely, but when one or more of the measures is partially or completely atomic the problem exhibits interesting and subtle behaviour. The key element is the order in which the norms are taken: The two norms,

$$\begin{aligned}\|f\|_{(L^p_\mu, L^q_\nu)} &= \left(\int\left(\int |f(x,y)|^p\, d\mu(x)\right)^{q/p} d\nu(y)\right)^{1/q} \quad \text{and}\\ \|f\|_{(L^q_\nu, L^p_\mu)} &= \left(\int\left(\int |f(x,y)|^q\, d\nu(y)\right)^{p/q} d\mu(x)\right)^{1/p}\end{aligned}$$

are related but far from identical. One relationship is Minkowski's integral inequality: If $p \le q$ then $\|f\|_{(L^p_\mu, L^q_\nu)} \le \|f\|_{(L^q_\nu, L^p_\mu)}$. For a multi-variable version of the Minkowski inequality, see [11]. The Kolmogorov-Nagumo theorem shows that they are far from identical. In fact these two norms can be equivalent only when $p = q$. (Or when one of the measures is just a finite number of atoms so that all norms are equivalent.) See [6] for references and for a proof of the Kolmogorov-Nagumo theorem in great generality.

We begin by reviewing some basic results for mixed norm Banach function spaces, in Section 2 and recalling the solution of the single variable inclusion problem for Lebesgue spaces, in Section 3. In Sections 4 and 5, we address the two-variable problem with the order of the norms reversed. The multivariable inclusion problem is considered in Section 6, both on its own and in situations where it reduces to the two-variable case.

Before continuing with Section 2 we introduce some definitions and notation. If (X, μ) is a σ-finite measure space, let L^+_μ denote the set of μ-measurable functions mapping X to $[0, \infty]$. Define L^p_μ to be the collection of all μ-measurable functions f mapping X to $\mathbb{R}$ (or $\mathbb{C}$) for which,

$$\|f\|_{p,\mu} \equiv \begin{cases} \left(\int |f|^p\, d\mu\right)^{1/p}, & 0 < p < \infty,\\ \operatorname{ess\,sup}_\mu |f|, & p = \infty,\end{cases}$$

is finite. When μ is counting measure we write ℓ^p and $\|f\|_p$, instead. A measurable subset E of X is an *atom* if $\mu(E) > 0$ and every measurable subset of E has measure $\mu(E)$ or 0.

For $p, q \in (0, \infty]$ define $p', p{:}q \in (-\infty, 0) \cup (0, \infty]$ by

$$\frac{1}{p'} + \frac{1}{p} = 1 \quad \text{and} \quad \frac{1}{p{:}q} = \frac{1}{q} - \frac{1}{p}.$$

Note that $1' = \infty$ and $p{:}p = \infty$. The p' notation is standard but $p{:}q$ is introduced here as a convenience to simplify complicated expressions appearing as exponents.

The non-increasing rearrangement $K^* = (K^*_1, K^*_2, \dots)$ of a (finite or infinite) sequence $K = (K_1, K_2, \dots)$ is defined in [4] as a function on $(0, \infty)$. However, it is readily seen to be a step function constant on the intervals $[n, n+1)$ for $n = 1, 2, \dots$, and may be identified with a sequence. We make this identification throughout.

2. Mixed Norm Banach Function Spaces

To define multivariable mixed norm Lebesgue spaces we extend the notation used in [3] and [11] so that we may keep track of exponents, measures, and spaces. It is simplest to state in terms of Banach function spaces. See [19] for relevant definitions. We assume that all our Banach function spaces have the Fatou property.

For $m = 1, \dots n$, let P_m be a Banach function space relative to a non-trivial σ-finite measure space (X_m, μ_m). Set $P = (P_1, \dots, P_n)$ and $\mu = \mu_1 \times \cdots \times \mu_n$. Let σ be a permutation of $\{1, \dots, n\}$ and write $\sigma(P) = (P_{\sigma(1)}, \dots, P_{\sigma(n)})$. For $f \in L^+_{\mu_1 \times \cdots \times \mu_n}$,

$$\|f\|_{\sigma(P)} = \Big\| \dots \Big\| \|f\|_{P_{\sigma(1)}} \Big\|_{P_{\sigma(2)}} \dots \Big\|_{P_{\sigma(n)}}.$$

The collection of all μ-measurable functions f for which $\|f\|_{\sigma(P)}$ is finite is denoted simply $\sigma(P)$, emphasizing that the mixed norm space depends only on the vector of single-variable norms. We stress that, regardless of the permutation σ, if $f = f(x_1, x_2, \dots, x_n)$, then the norm in the space P_m is always taken with respect to the variable x_m; with the other variables held fixed. It is important to point out that taking successive norms of a measurable function with respect to one or more variables always leaves us with a measurable function in the remaining variables. This is a consequence of Tonelli's theorem when the norms are Lebesgue norms with finite indices, but the Luxemburg-Gribanov theorem shows that it remains valid for Banach function norms satisfying the Fatou property. See [13]. (A counterexample in [14] shows that it is not enough to assume the weak Fatou property.) A version of the Luxemburg-Gribanov theorem for the more general mixed family-norms is in Theorem 5.1 of [18].

Proposition 2.1. *For $P = (P_1, \dots, P_n)$ and σ as above, $\sigma(P)$ is a Banach function space. Its associate space is $\sigma(P)' = \sigma(P_1', \dots, P_n')$.*

Proof. Both statements are given in [5]. The first is in a remark on page 158 and the second follows by simple induction from Theorem 3.12. □

To state the inclusion problem we need another mixed-norm space, in addition to $\sigma(P)$. Let $R = (R_1, \dots, R_n)$ be vector of Banach function spaces, relative to the non-trivial σ-finite measure spaces $(X_1, \kappa_1), \dots, (X_n, \kappa_n)$. Set $\kappa = \kappa_1 \times \cdots \times \kappa_n$ and let τ be a permutation of $\{1, \dots, n\}$. Note that the underlying sets $X_1, \dots, X_n$ are the same as for the spaces $P_1, \dots, P_n$, so the elements of $\sigma(P)$ and $\tau(R)$ are functions on the same domain. The inclusion problem asks whether or not $\sigma(P) \subseteq \tau(R)$. Since we are free to choose the "initial" order of the spaces $(X_1, \dots, X_n)$ there is no need for both permutations σ and τ. From now on we assume, without loss of generality, that τ is the identity permutation.

If, for any m, there is a μ_m-measurable set that is not κ_m-measurable, or a μ_m-null set that is not κ_m-null then inclusion fails. So we may assume that $\kappa_m \ll \mu_m$ for each m. On the other hand it is easy to see that inclusion holds if and only if it holds when, for each m, κ_m is restricted to the μ_m-measurable sets and μ_m is restricted to the support of the Radon-Nikodym derivative $[d\kappa_m/d\mu_m]$. By making these restrictions we may also assume that $\mu_m \ll \kappa_m$ for each m.

A standard argument shows that between Banach function spaces, set inclusion is equivalent to continuous inclusion. Accordingly, for each m we let C_m be the least constant, finite or infinite, such that

$$\|f\|_{R_m} \le C_m \|f\|_{P_m}, \quad f \in L^+_{\mu_m}. \tag{2.1}$$

Also, let C be the least constant, finite or infinite, such that

$$\|f\|_R \le C \|f\|_{\sigma(P)}, \quad f \in L^+_{\mu_1 \times \cdots \times \mu_n}. \tag{2.2}$$

Then C is finite if and only if $\sigma(P) \subseteq R$. Determining whether or not C is finite gives a qualitative answer to the inclusion problem. Finding the precise value of C gives a quantitive answer.

Theorem 2.2. *Suppose P, σ, and R are as above.*

(1) If $\sigma(P) \subseteq R$, then $P_m \subseteq R_m$ for each m and $C_1 \dots C_n \leq C$.
(2) If $P_m \subseteq R_m$ for each m then, $P \subseteq R$ and $C = C_1 \dots C_n$.

Proof. For the first statement, consider functions of the form $f(x_1, \dots, x_n) = f_1(x_1) \dots f_n(x_n)$, where $f_m \in L^+_{\mu_m}$. Both mixed norms split into a product of norms of $f_1, \dots, f_n$. Taking the supremum over all such $f_1, \dots, f_n$ gives the inequality. For the second statement, apply the single variable inclusion inequalities successively to get the reverse inequality in the "unpermuted case" when σ is the identity permutation. □

The first part of this theorem shows that all single-variable inclusions are necessary for a mixed-norm inclusion, regardless of the permutation σ. A similar result holds for certain two-variable inclusions, but these do depend on σ.

Lemma 2.3. *Suppose P, σ, and R are as above, and $\sigma(P) \subseteq R$. If $1 \leq i < j \leq n$ and $\sigma^{-1}(i) < \sigma^{-1}(j)$ then $(P_i, P_j) \subseteq (R_i, R_j)$. If $1 \leq i < j \leq n$ and $\sigma^{-1}(i) > \sigma^{-1}(j)$ then $(P_j, P_i) \subseteq (R_i, R_j)$.*

Proof. Suppose $1 \leq i < j \leq n$ and consider functions of the form

$$f(x_1, \dots, x_n) = g(x_i, x_j) \textstyle\prod_{m \neq i, m \neq j} f_m(x_m),$$

where $g \in L^+_{\mu_i \times \mu_j}$ and the f_m are fixed non-zero functions in $P_m \cap R_m$. Writing out the norm $\|f\|_R$ it is easy to see that $f \in R$ if and only if $g \in (R_i, R_j)$. The condition $\sigma^{-1}(i) < \sigma^{-1}(j)$ means that P_i appears to the left of P_j in $\sigma(P)$ and hence $f \in \sigma(P)$ if and only if $g \in (P_i, P_j)$. Now we have the implication,

$$g \in (P_i, P_j) \implies f \in \sigma(P) \implies f \in R \implies g \in (R_i, R_j).$$

The condition $\sigma^{-1}(i) > \sigma^{-1}(j)$ means that P_i appears to the right of P_j in P and hence $f \in \sigma(P)$ if and only if $g \in (P_j, P_i)$. In this case,

$$g \in (P_j, P_i) \implies f \in \sigma(P) \implies f \in R \implies g \in (R_i, R_j).$$

This completes the proof. □

Relative to the n-variable inclusion $\sigma(P) \subseteq R$ we refer to the inclusions given in the lemma as the *two-variable subinclusions.*

3. Single Variable Lebesgue Space Inclusions

To discuss the general Lebesgue space inclusion problem we need two Lebesgue spaces of functions on a single set. To solve the problem we have to understand the atomic structure of the underlying measures. We collect the necessary notation in,

$$\left\{ \begin{array}{c} p, r \in (0, \infty];\ 0 \neq \kappa \ll \mu \ll \kappa\ \sigma\text{-finite measures on } X;\ v = \left[\frac{d\kappa}{d\mu}\right]^{1/r}; \\ A = \sup\{\|gv\|_{r,\mu}/\|g\|_{p,\mu} : 0 \neq g \in L^+_\mu\}; \\ X = X_0 \dot\cup \big(\dot\cup_{i \in I} E_i\big);\ X_0 \text{ is atomless; } E_i \text{ is an atom for } i \in I; \\ k_i = \kappa(E_i);\ m_i = \mu(E_i);\ \text{and } M = (k_i^{1/r} m_i^{-1/p})_{i \in I}. \end{array} \right\} \tag{3.1}$$

As pointed out in Section 2, there is no loss of generality in assuming $\kappa \ll \mu \ll \kappa$. Since κ and μ are σ-finite, the Radon-Nikodym derivative exists and $v = [d\kappa/d\mu]^{1/r}$ is well-defined; when $r = \infty$ this is $v = \chi_{\{[d\kappa/d\mu]>0\}}$. Note that

$$\|g\|_{r,\kappa} = \|gv\|_{r,\mu}. \tag{3.2}$$

The supremum that defines A makes it the least constant, finite or infinite, in the inequality

$$\|gv\|_{r,\mu} \le A\|g\|_{p,\mu}, \quad g \in L^+_\mu. \tag{3.3}$$

Since κ and μ are a non-trivial σ-finite measures, $A > 0$. Observe that, as for Banach function spaces, the inclusion $L^p_\mu \subseteq L^r_\kappa$ holds if and only if A is finite. (When $0 < p < 1$, L^p_μ is not a Banach function space.)

Finally, since κ and μ are mutually absolutely continuous and σ-finite, a subset of X is an atom for κ if and only if it is an atom for μ. It is routine measure theory (see [7]) that X can be decomposed as a set X_0 that contains no atom and a collection of at most countably many disjoint atoms $E_i \subseteq X \setminus X_0$, for $i \in I$. Note every $f \in L^+_\mu$ is constant μ-almost everywhere on E_i. If $\mu(X \setminus X_0) = 0$ we say X is *atomless*, and if $\mu(X_0) = 0$ we say X is *purely atomic.*

Theorem 3.1. *Suppose (3.1). If $r \le p$, then $A = \|v\|_{p:r,\mu}$; in particular, $L^p_\mu \subseteq L^r_\kappa$ if and only if $v \in L^{p:r}_\mu$. If $r > p$, then $A = \infty$ unless X is purely atomic, in which case $A = \|M\|_\infty$; in particular, $L^p_\mu \subseteq L^r_\kappa$ if and only if X is purely atomic and $\{k_i^{1/r} m_i^{-1/p}, i \in I\}$ is bounded.*

Proof. If $r \le p$, the sharpness of Hölder's inequality, with indices p/r and $p{:}r/r$, shows that $A = \|v\|_{p:r,\mu}$.

Now suppose $r > p$. If X is not purely atomic then $\kappa(X_0) > 0$. For each $\varepsilon > 0$ small enough that $X_\varepsilon = \{x \in X_0 : [d\kappa/d\mu](x) > \varepsilon\}$ also has positive κ-measure, choose a subset E of X_ε having positive κ-measure less than $\varepsilon^{-2p:r/p}$. (Note that $p{:}r < 0$.) Since $E \subseteq X_\varepsilon$, $0 < \varepsilon\mu(E) \le \kappa(E)$. Taking $g = \chi_E$ in (3.3), and using (3.2), we find that $A \ge \kappa(E)^{1/r}\mu(E)^{-1/p} \ge \varepsilon^{-1/p} \to \infty$ as $\varepsilon \to 0$. Thus $A = \infty$.

Still in the case $r > p$, if X is purely atomic fix $g \in L^+_\mu$ and let g_i denote the value of g on E_i to get,

$$\|g\|^p_{r,\kappa} = \Big(\sum_{i\in I} g_i^r k_i\Big)^{p/r} \le \sum_{i\in I} g_i^p k_i^{p/r} \le \|M\|^p_\infty \sum_{i\in I} g_i^p m_i = \|M\|^p_\infty \|g\|^p_{p,\mu}.$$

Thus, $A \le \|M\|_\infty$. But with $g = \chi_{E_i}$, inequality (3.3) reduces to $A \ge k_i^{1/r} m_i^{-1/p}$ so $A \ge \sup_{i\in I} k_i^{1/r} m_i^{-1/p} = \|M\|_\infty$. □

4. Two Variable Mixed Norm Lebesgue Spaces

For two-variable inclusions we need two sets, with associated measures and indices. Rather than burden all the variables in (3.1) with the subscripts "1" and "2" we keep the originals and obtain a second bunch by moving up by one letter in two alphabets. We collect the notation in,

$$\left\{\begin{array}{c} q,s\in(0,\infty];\ 0\neq\lambda\ll\nu\ll\lambda\ \sigma\text{-finite measures on }Y;\ w=\left[\frac{d\lambda}{d\nu}\right]^{1/s};\\ B=\sup\{\|hw\|_{s,\nu}/\|h\|_{q,\nu}:0\neq h\in L^+_\nu\};\\ Y=Y_0\dot\cup\big(\dot\cup_{j\in J}F_j\big);\ Y_0\text{ is atomless};\ F_j\text{ are atoms for }j\in J;\\ l_j=\lambda(F_j);\ n_j=\nu(F_j);\ \text{and } N=(l_j^{1/s}n_j^{-1/q})_{j\in J}. \end{array}\right\} \tag{4.1}$$

By Proposition 2.2, both single variable inclusion problems are necessary conditions for the two variable inclusions, so there is no loss of generality in assuming that $\kappa\ll\mu\ll\kappa$ and $\lambda\ll\nu\ll\lambda$. We also have,

$$\|h\|_{s,\lambda}=\|hw\|_{s,\nu} \tag{4.2}$$

for all $h\in L^+_\nu$ and $B>0$ is the least constant, finite or infinite, in the inequality

$$\|hw\|_{s,\nu}\le B\|h\|_{q,\nu},\quad h\in L^+_\nu.$$

Part 2 of Proposition 2.2 shows that the two-variable “unpermuted” mixed norm inclusion problem $(L^p_\mu,L^q_\nu)\subseteq(L^r_\kappa,L^s_\lambda)$ reduces directly to the two single-variable inclusion problems $L^p_\mu\subseteq L^r_\kappa$ and $L^q_\nu\subseteq L^s_\lambda$. These one-variable inclusions were characterized in Theorem 3.1, so in this section we are free to focus only on the “permuted” inclusion problem,

$$(L^q_\nu,L^p_\mu)\subseteq(L^r_\kappa,L^s_\lambda). \tag{4.3}$$

It follows from (3.2) and (4.2), that

$$\big\|\|f\|_{r,\kappa}\big\|_{s,\lambda}=\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu}$$

for all $f\in L^+_{\mu\times\nu}$. Thus, for the quantitative version of the two variable inclusion problem (4.3), C is the least constant, finite or infinite, such that

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu}\le C\big\|\|f\|_{q,\mu}\big\|_{p,\nu},\quad f\in L^+_{\mu\times\nu}. \tag{4.4}$$

In view of Proposition 2.1, (L^q_ν,L^p_μ) and (L^r_κ,L^s_λ) are Banach function spaces provided all indices are at least 1, so (4.3) holds if and only if $C<\infty$. A peek ahead at Lemma 4.2 shows that this observation remains true for any positive indices.

Theorem 4.1. *Suppose (3.1), (4.1) and C is the best constant in (4.4).*

(1) Let $\max(q,s)\ge\min(p,r)$. Then $(L^q_\nu,L^p_\mu)\subseteq(L^r_\kappa,L^s_\lambda)$ if and only if $L^p_\mu\subseteq L^r_\kappa$ and $L^q_\nu\subseteq L^s_\lambda$. Moreover $C=AB$.

(2) Let $\max(q,s)<\min(p,r)$.

- *(a) If neither X nor Y is purely atomic then $C=\infty$, that is, the inclusion $(L^q_\nu,L^p_\mu)\subseteq(L^r_\kappa,L^s_\lambda)$ fails.*
- *(b) If X is purely atomic and Y is not purely atomic, then $(L^q_\nu,L^p_\mu)\subseteq(L^r_\kappa,L^s_\lambda)$ if and only if $L^q_\nu\subseteq L^s_\lambda$ and $M\in\ell^{p\dot{:}q}$. Moreover, if X is purely atomic and Y is atomless, then $C=B\|M\|_{p\dot{:}q}$.*
- *(c) If Y is purely atomic and X is not purely atomic, then $(L^q_\nu,L^p_\mu)\subseteq(L^r_\kappa,L^s_\lambda)$ if and only if $L^p_\mu\subseteq L^r_\kappa$ and $N\in\ell^{r\dot{:}s}$. Moreover, if Y is purely atomic and X is atomless, then $C=A\|N\|_{r\dot{:}s}$.*
- *(d) If X and Y are purely atomic, then,*
  - *(i) For $q\le s<p\le r$, $C=\|M^*N^*\|_{p\dot{:}s}$.*
  - *(ii) For $q\le s<r<p$, $C=GPP_{p\dot{:}s/p\dot{:}r}(N^{p\dot{:}s},M^{p\dot{:}r})^{1/p\dot{:}s}$, see Definition 5.6. In particular, $\|M^*N^*\|_{p\dot{:}s}\le C\le\min(A\|N\|_{r\dot{:}s},B\|M\|_{p\dot{:}s})$.*

(iii) For $s < q < p \le r$, $C = GPP_{p:s/q:s}(M^{p:s}, N^{q:s})^{1/p:s}$, see Definition 5.6. In particular, $\|M^*N^*\|_{p:s} \le C \le \min(A\|N\|_{p:s}, B\|M\|_{p:q})$.

(iv) For $s < q < r < p$, $\|M^*N^*\|_{p:s} \le C \le \min(A\|N\|_{r:s}, B\|M\|_{p:q})$.

Proof. Most of this section and the next will be devoted to proving these results. For (1) see Theorem 4.5. For (2a) see Theorem 4.7. For (2b) and (2c) see Theorems 4.8 and 4.9. For (2(d)i) see Theorem 5.5. For (2(d)ii) and (2(d)iii) see Theorems 5.8 and 5.9. Finally, for (2(d)iv) refer to Theorems 4.8, 4.9 and 5.3. □

We begin our analysis of two-variable inclusions by reducing the problem to the case that all Lebesgue indices lie in $(1, \infty)$. This will ensure that our Lebesgue spaces are Banach function spaces and will enable us to avoid considering special cases when one or more of the indices is equal to ∞.

Lemma 4.2. *Reduction by Substitution: Suppose (3.1), (4.1) and C is the best constant in (4.4). Fix $0 < t < \infty$ and let $\bar{p} = tp$, $\bar{q} = tq$, $\bar{r} = tr$, and $\bar{s} = ts$. Define $\bar{v}$, $\bar{A}$, $\bar{M}$, $\bar{w}$, $\bar{B}$, $\bar{N}$, and $\bar{C}$ as in (3.1), (4.1) and (4.4) but using $\bar{p}, \bar{q}, \bar{r}, \bar{s}$ in place of p, q, r, s. Then*

$$\bar{v} = v^{1/t},\ \bar{w} = w^{1/t},\ \bar{A} = A^{1/t},\ \bar{B} = B^{1/t},\ \bar{C} = C^{1/t},\ \bar{M}_i = M_i^{1/t}, \text{and } \bar{N}_j = N_j^{1/t}.$$

The conclusions of Theorem 4.1 are unchanged when all variables are replaced by their barred counterparts.

Proof. Routine verification is all that is required. □

Lemma 4.3. *Reduction by Duality: Suppose (3.1), (4.1) and C is the best constant in (4.4). Also suppose that $p, q, r, s \in [1, \infty]$. Let*

$$\bar{p} = s',\ \bar{q} = r',\ \bar{r} = q',\ \bar{s} = p',\ \text{and}$$

$$\bar{X} = Y,\ \bar{I} = J,\ \bar{\mu} = \nu,\ \bar{\kappa} = w^{q'}\nu,\ \bar{Y} = X,\ \bar{J} = I,\ \bar{\nu} = \mu,\ \bar{\lambda} = v^{p'}\mu.$$

Define $\bar{v}$, $\bar{A}$, $\bar{M}$, $\bar{w}$, $\bar{B}$, $\bar{N}$, and $\bar{C}$ as in (3.1), (4.1) and (4.4) but using the barred variables. Then,

$$\bar{v} = w,\ \bar{w} = v,\ \bar{A} = B,\ \bar{B} = A,\ \bar{C} = C,\ \bar{M} = N, \text{and } \bar{N} = M. \text{ Also,}$$

$$\bar{p}{:}\bar{q} = r{:}s,\ \bar{r}{:}\bar{s} = p{:}q,\ \bar{p}{:}\bar{s} = p{:}s,\ \bar{p}{:}\bar{r} = q{:}s, \text{and } \bar{q}{:}\bar{s} = p{:}r.$$

The conclusions of Theorem 4.1 are unchanged when all variables are replaced by their barred counterparts.

Proof. For $p, q, r, s \in [1, \infty]$, duality in Banach function spaces shows that constants A, B and C have equivalent definitions in terms of associate spaces. They are the least constants, finite or infinite, such that

$$\|gv\|_{p',\mu} \le A\|g\|_{r',\mu}, \quad g \in L^+_\mu; \tag{4.5}$$

$$\|hw\|_{q',\nu} \le B\|h\|_{s',\nu}, \quad h \in L^+_\nu; \quad \text{and} \tag{4.6}$$

$$\big\|\|fvw\|_{q',\nu}\big\|_{p',\mu} \le C\big\|\|f\|_{r',\mu}\big\|_{s',\nu}, \quad f \in L^+_{\mu\times\nu}; \tag{4.7}$$

respectively. But $\bar{v} = [d\bar{\kappa}/d\bar{\mu}]^{1/\bar{r}} = [w^{q'}d\nu/d\nu]^{1/q'} = w$ and, similarly, $\bar{w} = v$. Therefore, $\bar{A} = B$, $\bar{B} = A$, and $\bar{C} = C$.

For $j \in J$, let w_j denote the value of w takes ν-almost everywhere on F_j. Then $l_j = \lambda(F_j) = w_j^s n_j$ and $\bar{\kappa}(F_j) = w_j^{q'} n_j$. Therefore,

$$\bar{M}_j = \bar{\kappa}(F_j)^{1/\bar{r}}\bar{\mu}(F_j)^{-1/\bar{p}} = w_j n_j^{1/q:s} = l_j^{1/s} n_j^{-1/q} = N_j,$$

so $\bar{M} = N$. Similarly, $\bar{N} = M$. The last line of equations is trivial, included in the lemma for easy reference. Routine verification shows that overall the conclusions of Theorem 4.1 are unchanged, although several are interchanged. $\square$

Corollary 4.4. *To establish Theorem 4.1 for $p, q, r, s \in (0, \infty]$ it enough to prove it for $p, q, r, s \in (1, \infty)$.*

Proof. By Lemma 4.2, if Theorem 4.1 holds with indices p, q, r, s then it holds with indices tp, tq, tr, ts for any $t \in (0, \infty)$. Suppose it has been proved for $p, q, r, s \in (1, \infty)$. Then it is valid whenever $p, q, r, s \in [2, 4]$. Take $t = 1/2$ to see that it remains valid for $p, q, r, s \in [1, 2]$.

By Lemma 4.3, if Theorem 4.1 holds with indices p, q, r, s (for all spaces X and Y) then it holds with indices s', r', q', p' (for all spaces X and Y.) But we have it for all $p, q, r, s \in [1, 2]$ so it holds for all $p, q, r, s \in [2, \infty]$.

Another application of Lemma 4.2, shows that Theorem 4.1 holds whenever $p, q, r, s \in [2t, \infty]$ for any $t > 0$ so it holds whenever $p, q, r, s \in (0, \infty]$. $\square$

Minkowski's integral inequality is all that is needed to prove that the necessary condition provided by Proposition 2.2 is also sufficient for a large range of indices.

Theorem 4.5. *Suppose (3.1), (4.1) and C is the best constant in (4.4). If the indices satisfy $\min(p, r) \leq \max(q, s)$ then $C = AB$.*

Proof. There are four cases. Each one uses both of the single variable inclusions and one application of Minkowski's integral inequality. If $p \leq q$, then

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \leq A\big\|\|fw\|_{p,\mu}\big\|_{s,\nu} \leq AB\big\|\|f\|_{p,\mu}\big\|_{q,\nu} \leq AB\big\|\|f\|_{q,\nu}\big\|_{p,\mu}.$$

If $p \leq s$, then

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \leq A\big\|\|fw\|_{p,\mu}\big\|_{s,\nu} \leq A\big\|\|fw\|_{s,\nu}\big\|_{p,\mu} \leq AB\big\|\|f\|_{q,\nu}\big\|_{p,\mu}.$$

If $r \leq q$, then

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \leq B\big\|\|fv\|_{r,\mu}\big\|_{q,\nu} \leq B\big\|\|fv\|_{q,\nu}\big\|_{r,\mu} \leq AB\big\|\|f\|_{q,\nu}\big\|_{p,\mu}.$$

If $r \leq s$, then

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \leq \big\|\|fvw\|_{s,\nu}\big\|_{r,\mu} \leq A\big\|\|fw\|_{s,\nu}\big\|_{p,\mu} \leq AB\big\|\|f\|_{q,\nu}\big\|_{p,\mu}.$$

In each case we have $C \leq AB$. Part 1 of Proposition 2.2 gives the reverse inequality. $\square$

For the remainder of the section we consider only indices satisfying,

$$\max(q, s) < \min(p, r).$$

In this range it is insufficient to test inequality (4.4) only over factorable functions, those of the form $f(x, y) = g(x)h(y)$. Testing over this small class still gives the lower bound $AB \leq C$ but Theorem 4.7, below, presents a case in which AB may be finite but C is always infinite. This is a significant difference between the permuted case and the unpermuted case. In the latter, it suffices to test over factorable functions no matter what indices are involved.

We consider the more general class of block diagonal functions with factorable blocks, $f(x, y) = \sum_k g_k(x)h_k(y)$, where the g_k have disjoint supports in X, and the h_k have disjoint supports in Y. We will need to consider the norms on the individual blocks so we extend the definition of A and B as follows. If $X_k \subseteq X$ and $Y_k \subseteq Y$ let $A(X_k)$ and $B(Y_k)$ be the best constants in

$$\|gv\chi_{X_k}\|_{r,\mu} \leq A(X_k)\|g\chi_{X_k}\|_{p,\mu} \quad \text{and} \quad \|hw\chi_{Y_k}\|_{s,\nu} \leq B(Y_k)\|h\chi_{Y_k}\|_{q,\nu},$$

respectively, for all $g \in L^+_\mu$ and all $h \in L^+_\nu$. Clearly, $A(X) = A$, $A(\emptyset) = 0$, $B(Y) = B$, and $B(\emptyset) = 0$.

Theorem 4.6. *Suppose (3.1), (4.1), $\max(q,s) < \min(p,r)$, and C is the best constant in (4.4). If $X_1, X_2, \dots$ are disjoint subsets of X each having positive μ-measure and $Y_1, Y_2, \dots$ are disjoint subsets of Y each having positive ν-measure, then*

$$\Big(\sum_k A(X_k)^{p:s} B(Y_k)^{p:s}\Big)^{1/p:s}.$$

is the best constant in (4.4) when f is restricted to functions of the form $f(x,y) = \sum_k c_k g_k(x) h_k(y)$, where for each k, $g_k \in L^+_\mu$ is supported in X_k, $h_k \in L^+_\nu$ is supported in Y_k, and $c_k \geq 0$. In particular, it is a lower bound for C.

Proof. For f as above,

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} = \Big(\sum_k c_k^s \|g_k v\|_{r,\mu}^s \|h_k w\|_{s,\nu}^s\Big)^{1/s}$$

and

$$\big\|\|f\|_{q,\nu}\big\|_{p,\mu} = \Big(\sum_k c_k^p \|g_k\|_{p,\mu}^p \|h_k\|_{q,\nu}^p\Big)^{1/p}.$$

Fix any choice of non-zero g_k and h_k and let the coefficients c_k vary. Since $s < p$, the sharpness of Hölder's inequality implies that,

$$\sup\left\{\frac{\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu}}{\big\|\|f\|_{q,\nu}\big\|_{p,\mu}} : c_k \geq 0\right\} = \left(\sum_k \left(\frac{\|g_k v\|_{r,\mu}}{\|g_k\|_{p,\mu}}\right)^{p:s} \left(\frac{\|h_k w\|_{s,\nu}}{\|h_k\|_{q,\nu}}\right)^{p:s}\right)^{p:s}.$$

Taking the supremum of this expression over all g_k supported on X_k and over all h_k supported on Y_k gives,

$$\Big(\sum_k A(X_k)^{p:s} B(Y_k)^{p:s}\Big)^{p:s},$$

and completes the proof. □

This lower bound on C is infinite if the spaces X and Y both have non-trivial atomless parts.

Theorem 4.7. *Suppose (3.1), (4.1), $\max(q,s) < \min(p,r)$, and C is the best constant in (4.4). If neither X nor Y is purely atomic, then $C = \infty$.*

Proof. Since $\mu(X_0)$ and $\nu(Y_0)$ are non-zero and μ and ν are σ-finite, there exist $E \subseteq X_0$ and $F \subseteq Y_0$ such that $0 < \mu(E) < \infty$ and $0 < \nu(F) < \infty$. Set

$$E_\varepsilon = \{x \in E : [d\kappa/d\mu] > \varepsilon\} \quad \text{and} \quad F_\varepsilon = \{y \in F : [d\lambda/d\nu] > \varepsilon\},$$

where $\varepsilon > 0$ is chosen small enough so that $\mu(E_\varepsilon)$ and $\nu(F_\varepsilon)$ are both positive. If $C < \infty$ then $A < \infty$ and $B < \infty$ so, by Theorem 3.1, $r \leq p$ and $s \leq q$. Thus,

$$t \equiv \frac{p:s}{q:s} + \frac{p:s}{p:r} = \frac{\frac{1}{s} - \frac{1}{q} + \frac{1}{r} - \frac{1}{p}}{\frac{1}{s} - \frac{1}{p}} \in [0,1)$$

so there exist positive real numbers $\gamma_1, \gamma_2, \dots$, so that $\sum \gamma_k = 1$ but $\sum \gamma_k^t = \infty$. Partition E_ε into $X_1, X_2, \dots$ so that $\mu(X_k) = \gamma_k \mu(E_\varepsilon)$. Partition F_ε into $Y_1, Y_2, \dots$

so that $\nu(Y_k) = \gamma_k \nu(F_\varepsilon)$. Then,

$$A(X_k) \ge \frac{\|v\chi_{X_k}\|_{r,\mu}}{\|\chi_{X_k}\|_{p,\mu}} \ge \varepsilon\mu(X_k)^{1/r-1/p} = \varepsilon(\mu(E_\varepsilon)\gamma_k)^{1/p:r}$$

and

$$B(Y_k) \ge \frac{\|w\chi_{Y_k}\|_{s,\nu}}{\|\chi_{Y_k}\|_{q,\nu}} \ge \varepsilon\nu(Y_k)^{1/s-1/q} = \varepsilon(\nu(F_\varepsilon)\gamma_k)^{1/q:s}.$$

Thus, by Theorem 4.6,

$$C \ge \varepsilon^2 \mu(E_\varepsilon)^{1/p:r}\nu(F_\varepsilon)^{1/q:s}\Big(\sum \gamma_k^t\Big)^{1/p:s} = \infty.$$

This completes the proof. □

In view of the last result, we can restrict our attention to the case that at least one of X and Y is purely atomic.

Theorem 4.8. *Suppose (3.1), (4.1),* $\max(q,s) < \min(p,r)$*, and* C *is the best constant in (4.4). If* X *is purely atomic, then*

$$B(Y_0)\|M\|_{p:q} \le C \le B\|M\|_{p:q}. \tag{4.8}$$

If X *is purely atomic and* Y *is not purely atomic, then* $C < \infty$ *if and only if* $B\|M\|_{p:q} < \infty$*. If* X *is purely atomic and* Y *is atomless then* $C = B\|M\|_{p:q}$*.*

Proof. By Corollary 4.4 we may suppose $p,q,r,s \in (1,\infty)$. Let $f \in L^+_{\mu\times\nu}$. For ν-almost every y the function $x \mapsto f(x,y)$ is μ-measurable and hence constant μ-almost everywhere on each atom of μ. Let $f_i(y)$ denote the value of this function on E_i. Then,

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \le B\big\|\|fv\|_{r,\mu}\big\|_{q,\nu} = B\Big\|\Big(\sum_{i\in I} f_i^r k_i\Big)^{1/r}\Big\|_{q,\nu}.$$

But $q < r$, so this is no greater than

$$B\Big\|\Big(\sum_{i\in I} f_i^q k_i^{q/r}\Big)^{1/q}\Big\|_{q,\nu} = B\Big(\sum_{i\in I}\|f_i\|_{q,\nu}^q k_i^{q/r}\Big)^{1/q},$$

and Hölder's inequality completes the proof of the upper bound on C in (4.8), giving,

$$\big\|\|fvw\|_{r,\mu}\big\|_{s,\nu} \le B\|M\|_{p:q}\Big(\sum_{i\in I}\|f_i\|_{q,\nu}^p m_i\Big)^{1/p} = B\|M\|_{p:q}\big\|\|f\|_{q,\nu}\big\|_{p,\mu}.$$

For the lower bound there is nothing to prove if $C = \infty$ or if Y is purely atomic, so suppose $C < \infty$ and $\nu(Y_0) > 0$. Since $AB \le C$ it follows that $B < \infty$. By Theorem 3.1, the finiteness of B when Y is not purely atomic can occur only when $s \le q$. We set up to apply Theorem 4.6 by taking $X_i = E_i$. Clearly, $A(X_i) = k_i^{1/r} m_i^{-1/p}$ and hence,

$$\sum_{i\in I} A(X_i)^{p:q} = \|M\|_{p:q}^{p:q}. \tag{4.9}$$

To choose Y_i we consider two cases. If $s < q$ then, by Theorem 3.1,

$$0 \le \int_{Y_0} w^{q:s}\, d\nu = B(Y_0)^{q:s} \le B^{q:s} < \infty$$

so if I_0 is any finite subset of I, we may partition Y_0 into subsets Y_i, $i \in I_0$ so that

$$B(Y_i)^{q:s} = \int_{Y_i} w^{q:s}\, d\nu = B(Y_0)^{q:s} \frac{A(X_i)^{p:q}}{\sum_{\eta \in I_0} A(X_\eta)^{p:q}}.$$

Theorem 4.6 implies,

$$C \geq \Big(\sum_{i \in I_0} (A(X_i)B(Y_i))^{p:s} \Big)^{1/p:s} = B(Y_0) \Big(\sum_{i \in I_0} A(X_i)^{p:q} \Big)^{1/p:q}.$$

Taking the supremum of the right-hand side over all finite subsets of I, and using (4.9), gives the lower bound.

In the remaining case, $s = q$. We can apply Theorem 3.1 again to obtain, $\operatorname{ess\,sup}_\nu w\chi_{Y_0} = B(Y_0) \leq B < \infty$ so for any $\varepsilon > 0$ we may partition the set $\{y \in Y_0 : w(y) > B(Y_0) - \varepsilon\}$ into sets $Y_1, Y_2, \ldots$ of positive ν-measure so that $B(Y_i) \geq B(Y_0) - \varepsilon$ for each i. Theorem 4.6 and (4.9) imply,

$$C \geq \left(\sum_{i \in I} (A(X_i)B(Y_i))^{p:q} \right)^{1/p:q} \geq (B(Y_0) - \varepsilon)\|M\|_{p:q}$$

Let $\varepsilon \to 0$ to get the the lower bound. This completes the proof of (4.8).

For second statement of the theorem, note that if Y is not purely atomic then $B(Y_0) > 0$. The upper bound for C in (4.8) shows that $C < \infty$ whenever $B\|M\|_{p:q} < \infty$. On the other hand, if $C < \infty$ and $B(Y_0) > 0$ in (4.8), then $\|M\|_{p:q} < \infty$. But Proposition 2.2 shows that $C < \infty$ implies $B < \infty$. Thus, $B\|M\|_{p:q} < \infty$. For the last statement, just observe that if Y is atomless then $Y = Y_0$ so $B(Y_0) = B$. □

There is a corresponding theorem for the case when Y is purely atomic. Because we are in the permuted case, the inclusion problem is not symmetric in X and Y. However, duality enables us to deduce this result from the previous theorem.

Theorem 4.9. *Suppose (3.1), (4.1),* $\max(q, s) < \min(p, r)$, *and C is the best constant in (4.4). If Y is purely atomic, then*

$$A(X_0)\|N\|_{r:s} \leq C \leq A\|N\|_{r:s}.$$

If Y is purely atomic and X is not purely atomic, then $C < \infty$ if and only if $A\|N\|_{r:s} < \infty$. If Y is purely atomic and X is atomless then $C = A\|N\|_{r:s}$.

Proof. In addition to many other changes, the duality transformation of Lemma 4.3 interchanges X and Y and takes $A(X_0)$ to $B(Y_0)$. So if Y is purely atomic, Theorem 4.8 may be applied to the transformed problem. Its conclusions, when translated using Lemma 4.3, establish the present theorem. □

The last two theorems give a complete qualitative answer to the continuous inclusion question when at least one measure is not purely atomic. Determination of the exact value of C remains open in the case that one measure is purely atomic and the other is neither purely atomic nor atomless. A formula for the best constant in this case would give a formula for the best constant in the case when both spaces are purely atomic. This is unlikely; see Definition 5.6 and Theorems 5.8 and 5.9.

5. The Two Variable Case: Purely Atomic Measures

When both spaces are purely atomic the two-variable inclusion problem presents much more suble behavior than in the cases covered in the previous section. Using the notation of (3.1) and (4.1), we may rewrite (4.4) so that C is the least constant, finite or infinite, such that

$$\Big(\sum_{j\in J}\Big(\sum_{i\in I} f_{i,j}^r k_i\Big)^{s/r} l_j\Big)^{1/s} \le C\Big(\sum_{i\in I}\Big(\sum_{j\in J} f_{i,j}^q n_j\Big)^{p/q} m_i\Big)^{1/p}, \tag{5.1}$$

whenever $f_{i,j}\ge 0$ for $i\in I$ and $j\in J$. Here $f_{i,j}$ is just the value that the measurable function f takes on the atom $E_i\times F_j$ of $X\times Y$. It is convenient, in the case of purely atomic spaces, to identify functions on $X\times Y$ with functions on $I\times J$.

To determine the least value of C above it is enough, in view of the monotone convergence theorem, to consider functions f of finite support in $I\times J$. And, by homogeneity, it is enough to restrict these to the unit ball in (L^q_ν, L^p_μ). The constant C^s is the supremum of

$$\max\Big\{\sum_{j\in J_0}\Big(\sum_{i\in I_0} f_{i,j}^r k_i\Big)^{s/r} l_j : f_{i,j}\ge 0, \sum_{i\in I_0}\Big(\sum_{j\in J_0} f_{i,j}^q n_j\Big)^{p/q} m_i = 1\Big\}, \tag{5.2}$$

taken over all finite subsets $I_0\subseteq I$ and $J_0\subseteq J$. The maximum exists because it is supremum of a continuous function over a closed and bounded subset of a finite-dimensional real vector space. (Values of f off $I_0\times J_0$ are irrelevant and may be assumed to be zero.)

We begin our analysis of this case with a bit of calculus that will, for certain index ranges, restrict the functions at which the maximum in (5.2) may occur.

Theorem 5.1. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $1<q\le s<\min(p,r)<\infty$. Fix finite subsets $I_0\subseteq I$ and $J_0\subseteq J$. Let $f\ge 0$ be a function that has the fewest possible non-zero entries among those for which the maximum in (5.2) occurs. Then for any $\eta\in I_0$ and any distinct $\varphi,\psi\in J_0$, at most one of $f_{\eta,\varphi}$ and $f_{\eta,\psi}$ is non-zero.*

Proof. Suppose both $f_{\eta,\varphi}$ and $f_{\eta,\psi}$ are strictly positive. Then

$$a = (f_{\eta,\varphi}^q + \theta/n_\varphi)^{1/q} \quad\text{and}\quad b = (f_{\eta,\psi}^q - \theta/n_\psi)^{1/q}$$

are defined as functions of θ in the closed neighbourhood $[-f_{\eta,\varphi}^q n_\varphi, f_{\eta,\psi}^q n_\psi]$ of zero. Note that $a(0)=f_{\eta,\varphi}$, $b(0)=f_{\eta,\psi}$ and $a^q n_\varphi + b^q n_\psi$ is constant. Additionally, $a'>0$ and $b'<0$ on $(-f_{\eta,\varphi}^q n_\varphi, f_{\eta,\psi}^q n_\psi)$. Working on this open interval, we differentiate $a^q n_\varphi + b^q n_\psi$ to get,

$$a^{q-1}a' n_\varphi = b^{q-1}(-b')n_\psi.$$

Taking the logarithm of both sides and differentiating again shows that

$$(q-1)\frac{a'}{a} + \frac{a''}{a'} = (q-1)\frac{b'}{b} + \frac{b''}{b'}. \tag{5.3}$$

Define $\bar f$ to agree with f except that $\bar f_{\eta,\varphi}=a$ and $\bar f_{\eta,\psi}=b$. Then $\bar f = f$ when $\theta=0$ and, since $a^q n_\varphi + b^q n_\psi$ is constant,

$$\sum_{i\in I_0}\Big(\sum_{j\in J_0}\bar f_{i,j}^q n_j\Big)^{p/q} m_i = 1$$

for all θ. Thus, for each θ, $\bar f$ is among the functions over which the maximum is taken in (5.2), and so

$$\sum_{j\in J_0}\Big(\sum_{i\in I_0}\bar f_{i,j}^r k_i\Big)^{s/r} l_j \tag{5.4}$$

has a maximum at $\theta = 0$. But only two terms in (5.4) vary with θ so, setting

$$A = \sum_{i \in I_0} f_{i,\varphi}^r k_i \quad \text{and} \quad B = \sum_{i \in I_0} f_{i,\psi}^r k_i,$$

we see that

$$(A - f_{\eta,\varphi}^r k_\eta + a^r k_\eta)^{s/r} l_\varphi + (B - f_{\eta,\psi}^r k_\eta + b^r k_\eta)^{s/r} l_\psi \tag{5.5}$$

has a maximum at $\theta = 0$. Therefore, at $\theta = 0$, its derivative is zero and its second derivative is at most zero. The first derivative is,

$$s(A - f_{\eta,\varphi}^r k_\eta + a^r k_\eta)^{(s-r)/r} a^{r-1} a' k_\eta l_\varphi - s(B - f_{\eta,\psi}^r k_\eta + b^r k_\eta)^{(s-r)/r} b^{r-1} (-b') k_\eta l_\psi.$$

This is a difference of strictly positive functions x and y such that $x(0) = y(0)$ and $x'(0) \leq y'(0)$. It follows that $(\log x)'(0) = x'(0)/x(0) \leq y'(0)/y(0) = (\log y)'(0)$. Therefore,

$$(s-r)\frac{a(0)^r}{A}\frac{a'(0)}{a(0)}k_\eta + (r-1)\frac{a'(0)}{a(0)} + \frac{a''(0)}{a'(0)} \leq (s-r)\frac{b(0)^r}{B}\frac{b'(0)}{b(0)}k_\eta + (r-1)\frac{b'(0)}{b(0)} + \frac{b''(0)}{b'(0)}.$$

By (5.3) this simplifies to

$$(s-r)\frac{a(0)^r}{A}\frac{a'(0)}{a(0)}k_\eta + (r-q)\frac{a'(0)}{a(0)} \leq (s-r)\frac{b(0)^r}{B}\frac{b'(0)}{b(0)}k_\eta + (r-q)\frac{b'(0)}{b(0)},$$

and dividing both sides by the $a(0)^{q-1}a'(0)n_\varphi = b(0)^{q-1}(-b'(0))n_\psi > 0$ we get

$$\frac{(s-q)a(0)^r k_\eta + (r-q)(A - a(0)^r k_\eta)}{Aa(0)^q n_\varphi} \leq -\frac{(s-q)b(0)^r k_\eta + (r-q)(B - b(0)^r k_\eta)}{Bb(0)^q n_\psi}.$$

But $0 < a(0)^r k_\eta = f_{\eta,\varphi}^r k_\eta \leq A$ and $0 < b(0)^r k_\eta = f_{\eta,\psi}^r k_\eta \leq B$ so the left-hand side is at least zero and the right-hand side is at most zero. Thus, both sides are zero and so $s = q$, $A = f_{\eta,\varphi}^r k_\eta$ and $B = f_{\eta,\psi}^r k_\eta$. Using these equations, and recalling the definitions of a and b, we find that (5.5) becomes

$$a^q k_\eta^{q/r} l_\varphi + b^q k_\eta^{q/r} l_\psi = k_\eta^{q/r}(f_{\eta,\varphi}^q l_\varphi + f_{\eta,\psi}^q l_\psi) + k_\eta^{q/r}(l_\varphi/n_\varphi - l_\psi/n_\psi)\theta.$$

This straight line has a maximum at $\theta = 0$ and is therefore constant. It follows that (5.4) is also constant so the maximum in (5.2) occurs at $\bar{f}$ for each $\theta \in (-f_{\eta,\varphi}^q n_\varphi, f_{\eta,\psi}^q n_\psi)$ and, by continuity, for each $\theta \in [-f_{\eta,\varphi}^q n_\varphi, f_{\eta,\psi}^q n_\psi]$. But when θ is one of the endpoints of this interval, $\bar{f}$ has one fewer non-zero entry than f. This contradicts the choice of f and completes the proof. □

A similar argument proves the following.

Theorem 5.2. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $1 < \max(q,s) < p \leq r < \infty$. Fix finite subsets $I_0 \subseteq I$ and $J_0 \subseteq J$. Let $f \geq 0$ be a function that has the fewest possible non-zero entries among those for which the maximum in (5.2) occurs. Then for any distinct $\eta, \xi \in I_0$ and any $\varphi \in J_0$, at most one of $f_{\eta,\varphi}$ and $f_{\xi,\varphi}$ is non-zero.*

By combining the last two theorems we can give a quantitative answer to the inclusion problem in the case $q \leq s < p \leq r$. But first we need another application of Theorem 4.6.

Theorem 5.3. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $\max(q,s) < \min(p,r)$. Then $C \geq \|M^* N^*\|_{p:s}$.*

Proof. If $i_1,\dots,i_K$ are distinct elements of I and $j_1,\dots,j_K$ are distinct elements of J then Theorem 4.6 shows that,

$$C \ge \Big(\sum_{k=1}^K A(E_{i_k})^{p:s}B(F_{j_k})^{p:s}\Big)^{1/p:s} = \Big(\sum_{k=1}^K M_{i_k}^{p:s}N_{j_k}^{p:s}\Big)^{1/p:s}.$$

Basic properties of rearrangements, see 2.1.7, 2.2.3, and 2.2.7 of [4], show that taking the supremum over all choices of $i_1,\dots,i_K$ and $j_1,\dots,j_K$ for all K yields,

$$C \ge \Big(\sum_{k=1}^\infty (M_k^*)^{p:s}(N_k^*)^{p:s}\Big)^{1/p:s}.$$

This completes the proof. $\square$

Corollary 5.4. *Suppose X and Y are infinite sets with counting measure. The inclusion $(\ell^q(Y),\ell^p(X)) \subseteq (\ell^r(X),\ell^s(Y))$ holds if and only if $p\le r$, $q\le s$, and $p\le s$. In this case, the norm of the inclusion is* 1.

Proof. By Theorem 3.1, the first two index conditions imply the inclusions $\ell^p(X)\subseteq \ell^r(X)$ and $\ell^q(Y)\subseteq\ell^s(Y)$, both with norm 1. Using the third condition, Theorem 4.5 shows that $(\ell^q(Y),\ell^p(X)) \subseteq (\ell^r(X),\ell^s(Y))$, again with norm 1.

For the converse, the necessity of the single-variable inclusions, together with Theorem 3.1, implies $p\le r$ and $q\le s$. If $s<p$ then Theorem 5.3 shows $C=\infty$ because $M=N=(1,1,\dots)$. This contradiction proves that $p\le s$ and completes the proof. $\square$

In the range of indices where both Theorems 5.1 and 5.2 apply, the lower bound just established is also an upper bound.

Theorem 5.5. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $q\le s<p\le r$. Then $C=\|M^*N^*\|_{p:s}$.*

Proof. First suppose $1<q\le s<p\le r<\infty$. To get an upper bound for C, Theorems 5.1 and 5.2 show it is sufficient to test (5.1) over functions supported on a set $H\subseteq I\times J$ such that for $(i,j),(\bar i,\bar j)\in H$, $(i,j)\ne(\bar i,\bar j)$ implies $i\ne\bar i$ and $j\ne\bar j$. Let $\mathcal{H}$ be the collection of all such sets H. If f is supported on $H\in\mathcal{H}$ then (5.1) reduces to

$$\Big(\sum_{(i,j)\in H} f_{i,j}^s k_i^{s/r} l_j\Big)^{1/s} \le C\Big(\sum_{(i,j)\in H} f_{i,j}^p n_j^{p/q} m_i\Big)^{1/p}.$$

The sharpness of Hölder's inequality yields,

$$\sup_{f|_H=f}\frac{\Big(\sum_{(i,j)\in H} f_{i,j}^s k_i^{s/r} l_j\Big)^{1/s}}{\Big(\sum_{(i,j)\in H} f_{i,j}^p n_j^{p/q} m_i\Big)^{1/p}} = \Big(\sum_{(i,j)\in H} M_i^{p:s}N_j^{p:s}\Big)^{1/p:s}.$$

Thus, C is equal to the supremum over all $H\in\mathcal{H}$ of this last quantity. But the Hardy-Littlewood-Pólya inequality, Theorem 2.2.2 in [4], shows that

$$C=\sup_{H\in\mathcal{H}}\Big(\sum_{(i,j)\in H} M_i^{p:s}N_j^{p:s}\Big)^{1/p:s} \le \Big(\sum_{k=1}^\infty (M_k^*)^{p:s}(N_k^*)^{p:s}\Big)^{1/p:s} = \|M^*N^*\|_{p:s}.$$

Theorem 5.3 provides the reverse inequality. This completes the proof in the case $1<q\le s<p\le r<\infty$. By Corollary 4.4, it holds for $0<q\le s<p\le r\le\infty$. $\square$

When $s < q < p \le r$ or $q \le s < r < p$ we can apply one, but not both, of Theorems 5.1 and 5.2. This allows us to express C as a supremum over a restricted set of functions and this supremum can be abstracted as the solution to the following optimization problem.

Definition 5.6. *A Generalized Partition Problem: Given $\gamma \in (0,1)$ and two non-trivial non-negative sequences a_i, $i \in I$ and b_j, $j \in J$ let*

$$GPP_\gamma(a,b) = \sup \sum_{i\in I} a_i \Big(\sum_{j\in J(i)} b_j \Big)^\gamma,$$

where the supremum is taken over all partitions $\{J(i) : i \in I\}$ of J.

We don't propose to give a solution to this problem. Indeed, the simple special case $I = \{1,2\}$ with $a_1 = a_2 = 1$, and $J = \{1,\dots,N\}$ with $b_1,\dots,b_N$ positive integers, reduces to the optimization version of the Partition Problem, known to be NP-Hard. We content ourselves with a few observations and one example.

(1) If $GPP_\gamma(a,b) < \infty$ then $b \in \ell^1$ and $a \in \ell^\infty$. To see this, fix η and consider the trivial partition in which $J(\eta) = J$ and $J(i) = \emptyset$ for $i \ne \eta$. Then

$$GPP_\gamma(a,b) \ge a_\eta \Big(\sum_{j\in J} b_j \Big)^\gamma$$

and, taking the supremum over all η,

$$GPP_\gamma(a,b) \ge \|a\|_\infty \|b\|_1^\gamma.$$

(2) If $b \in \ell^1$ and $a \in \ell^{1/(1-\gamma)}$ then $GPP_\gamma(a,b) < \infty$. For this, apply Hölder's inequality with indices $1/\gamma$ and $1/(1-\gamma)$ to get

$$\sum_{i\in I} a_i \Big(\sum_{j\in J(i)} b_j \Big)^\gamma \le \|a\|_{1/(1-\gamma)} \Big(\sum_{i\in I} \sum_{j\in J(i)} b_j \Big)^\gamma = \|a\|_{1/(1-\gamma)} \|b\|_1^\gamma.$$

Take the supremum over all partitions to get $GPP_\gamma(a,b) \le \|a\|_{1/(1-\gamma)} \|b\|_1^\gamma$.

(3) If $a \in \ell^\infty$ and $b \in \ell^\gamma$ then $GPP_\gamma(a,b) < \infty$. Replace each a_i by its upper bound to get,

$$\sum_{i\in I} a_i \Big(\sum_{j\in J(i)} b_j \Big)^\gamma \le \|a\|_\infty \sum_{i\in I} \sum_{j\in J(i)} b_j^\gamma = \|a\|_\infty \|b\|_\gamma^\gamma.$$

Take the supremum over all partitions to get $GPP_\gamma(a,b) \le \|a\|_\infty \|b\|_\gamma^\gamma$.

(4) If $GPP_\gamma(a,b) < \infty$ then $a_i^*(b_i^*)^\gamma \in \ell^1$. Consider only those trivial partitions for which each $J(i)$ is a singleton. As in Theorem 5.5 one obtains,

$$GPP_\gamma(a,b) \ge \sum_{i=1}^\infty a_i^*(b_i^*)^\gamma = \|a^*(b^*)^\gamma\|_1.$$

Example 5.7. *For any $\gamma \in (0,1)$ there exist positive, non-increasing sequences a and b for which $GPP_\gamma(a,b) = \infty$, but the supremum over the trivial partitions from Observations (1) and (4) above gives a finite value.*

Proof. Fix $\gamma \in (0,1)$ and let Q be the least integer greater than or equal to $1/(1-\gamma)$. Let $x_k = k^{-2/\gamma}$, $m_k = 2^k$ and $n_k = 2^{Qk^2}$, for $k = 1,2,\dots$. (For convenience, set $m_0 = n_0 = 0$.) Define

$$a_j = \sum_{k=1}^\infty \frac{(m_k n_k)^\gamma}{n_k} \chi_{(0,n_k]}(j) \quad \text{and} \quad b_j = \sum_{k=1}^\infty \frac{x_k}{m_k n_k} \chi_{(0,m_k n_k]}(j).$$

Clearly, a and b are non-negative, non-increasing sequences. For each integer $k \geq 1$, define a partition of $\{1, 2, \dots\}$ by setting $J(i) = \{(i-1)m_k + 1, \dots, im_k\}$ for $i = 1, \dots, n_k$, $J(n_k + 1) = \{n_k m_k + 1, \dots\}$, and $J(i) = \emptyset$ for $i > n_k + 1$. Then

$$GPP_\gamma(a,b) \geq \sum_{i=1}^{n_k} a_i \Big(\sum_{j \in J(i)} b_j \Big)^\gamma \geq \sum_{i=1}^{n_k} \frac{(m_k n_k)^\gamma}{n_k} \Big(\sum_{j=(i-1)m_k+1}^{im_k} \frac{x_k}{m_k n_k} \Big)^\gamma = (x_k m_k)^\gamma,$$

which is unbounded as $k \to \infty$. Thus $GPP_\gamma(a,b) = \infty$.

For $k \geq 1$, $\gamma k - (1-\gamma)Qk^2 \leq k - k^2 \leq 1 - k$, so

$$\|a\|_\infty = a_1 = \sum_{k=1}^\infty \frac{(m_k n_k)^\gamma}{n_k} = \sum_{k=1}^\infty 2^{\gamma k - (1-\gamma)Qk^2} \leq 2.$$

Also,

$$\|b\|_1 = \sum_{k=1}^\infty \frac{x_k}{m_k n_k} \sum_{j=1}^\infty \chi_{(0, m_k n_k]}(j) = \sum_{k=1}^\infty \frac{x_k}{m_k n_k} m_k n_k = \sum_{k=1}^\infty x_k < \infty.$$

Therefore, $\|a\|_\infty \|b\|_1^\gamma < \infty$. Thus, the partitions from Observation (1) are not enough to show that $GPP_\gamma(a,b) = \infty$.

To see that the partitions used to get Observation (4) are not enough to show that $GPP_\gamma(a,b) = \infty$ either, we need to show that

$$\sum_{j=1}^\infty a_j b_j^\gamma < \infty.$$

(Since a and b are non-negative and non-increasing, they coincide with their rearrangements.) Let K be a positive integer and observe that $\gamma - (1-\gamma)Q(k+K) \leq -1$ for all $k \geq 1$. If $n_{K-1} < j \leq n_K$, then

$$a_j = \sum_{k=K}^\infty \frac{(m_k n_k)^\gamma}{n_k} = \frac{(m_K n_K)^\gamma}{n_K} \sum_{k=K}^\infty 2^{(k-K)(\gamma - (1-\gamma)Q(k+K))} \leq 2 \frac{(m_K n_K)^\gamma}{n_K}.$$

Also, if $m_{K-1} n_{K-1} < j \leq m_K n_K$, then

$$b_j = \sum_{k=K}^\infty \frac{x_k}{m_k n_k} \leq \frac{x_K}{m_K n_K} \sum_{k=K}^\infty \frac{m_K}{m_k} = \frac{x_K}{m_K n_K} \sum_{k=K}^\infty 2^{K-k} = 2 \frac{x_K}{m_K n_K}.$$

It is easy to check that $n_{k-1} \leq m_{k-1} n_{k-1} < n_k < m_k n_k < n_{k+1}$ for each positive integer k. So our estimates of a_j and b_j give

$$\sum_{j=1+m_{k-1}n_{k-1}}^{n_k} a_j b_j^\gamma \leq 2 \frac{(m_k n_k)^\gamma}{n_k} \left(2 \frac{x_k}{m_k n_k} \right)^\gamma (n_k - m_{k-1} n_{k-1}) \leq 2^{1+\gamma} x_k^\gamma,$$

and, since $k + \gamma - (1-\gamma)Q(2k+1) \leq 0$,

$$\sum_{j=1+n_k}^{m_k n_k} a_j b_j^\gamma \leq 2 \frac{(m_{k+1} n_{k+1})^\gamma}{n_{k+1}} \left(2 \frac{x_k}{m_k n_k} \right)^\gamma m_k n_k \leq 2^{1+\gamma} x_k^\gamma.$$

Now we have,

$$\sum_{j=1}^\infty a_j b_j^\gamma = \sum_{k=1}^\infty \left(\sum_{j=1+m_{k-1}n_{k-1}}^{n_k} a_j b_j^\gamma + \sum_{j=1+n_k}^{m_k n_k} a_j b_j^\gamma \right) \leq 2^{2+\gamma} \sum_{k=1}^\infty x_k^\gamma < \infty.$$

$\square$

The next two theorems give quantitative answers to the inclusion problem in the cases $q \le s < r < p$ and $s < q < p \le r$. However, these answers are given in terms of solutions of the Generalized Partition Problem of Definition 5.6. They should be viewed as indications of the intractibility of the inclusion problem in these cases rather than solutions of the problem. On the other hand, Observations (1)-(4) give some necessary conditions and some sufficient conditions for the finiteness of GPP. These conditions can be recast as conditions for the inclusion problem, giving easy to verify conditions that imply or are implied by the inclusion $(L^q_\nu, L^p_\mu) \subseteq (L^r_\kappa, L^s_\lambda)$.

Theorem 5.8. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $q \le s < r < p$. Then $C = GPP_{p:s/p:r}(N^{p:s}, M^{p:r})^{1/p:s}$. Consequently,*

$$\|M^*N^*\|_{p:s} \le C \le \min(A\|N\|_{r:s}, B\|M\|_{p:s}).$$

Proof. First suppose $1 < q \le s < r < p < \infty$. Theorem 5.1 shows that for this range of indices it is sufficient to test inequality (5.1) over functions supported on a set $H \subseteq I \times J$ such that for $(i,j), (\bar{i}, \bar{j}) \in H$, $(i,j) \ne (\bar{i}, \bar{j})$ implies $i \ne \bar{i}$. Let $\mathcal{H}$ be the collection of all such sets H. For each $H \in \mathcal{H}$, let $H(j) = \{i : (i,j) \in H\}$ and note that the sets $H(j)$ are disjoint. If f is supported on $H \in \mathcal{H}$ then (5.1) reduces to

$$\Big(\sum_{j\in J}\Big(\sum_{i\in H(j)} f^r_{i,j}k_i\Big)^{s/r} l_j\Big)^{1/s} \le C\Big(\sum_{j\in J}\sum_{i\in H(j)} f^p_{i,j} m_i n_j^{p/q}\Big)^{1/p}.$$

Since the sets $H(j)$ are disjoint and the $f_{i,j}$ are arbitrary, we are free to replace $f_{i,j}$ by $c_j f_{i,j}$ for arbitrary $c_j \ge 0$. This gives,

$$\begin{aligned}
&\sup_{f|_H=f} \frac{\Big(\sum_{j\in J}\Big(\sum_{i\in H(j)} f^r_{i,j}k_i\Big)^{s/r} l_j\Big)^{1/s}}{\Big(\sum_{j\in J}\sum_{i\in H(j)} f^p_{i,j} m_i n_j^{p/q}\Big)^{1/p}}\\
&= \sup_{f|_H=f}\sup_{c_j\ge 0} \frac{\Big(\sum_{j\in J} c_j^s\Big(\sum_{i\in H(j)} f^r_{i,j}k_i\Big)^{s/r} l_j\Big)^{1/s}}{\Big(\sum_{j\in J} c_j^p\Big(\sum_{i\in H(j)} f^p_{i,j} m_i\Big) n_j^{p/q}\Big)^{1/p}}\\
&= \sup_{f|_H=f}\Big(\sum_{j\in J}\Big(\Big(\sum_{i\in H(j)} f^r_{i,j}k_i\Big)^{1/r}\Big(\sum_{i\in H(j)} f^p_{i,j}m_i\Big)^{-1/p}\ell_j^{1/s} n_j^{-1/q}\Big)^{p:s}\Big)^{1/p:s}\\
&= \Big(\sum_{j\in J} N_j^{p:s}\Big(\sum_{i\in H(j)} M_i^{p:r}\Big)^{p:s/p:r}\Big)^{1/p:s}.
\end{aligned}$$

Taking the supremum of the last expression over all sets $H \in \mathcal{H}$ shows that

$$C = GPP_{p:s/p:r}(N^{p:s}, M^{p:r})^{1/p:s}.$$

Note that $0 < p{:}s/p{:}r < 1$.

The lower bound $\|M^*N^*\|_{p:s} \le C$ was proved in Theorem 5.3. (It also follows from Observation (4).) With $\gamma = p{:}s/p{:}r$, $a = N^{p:s}$, and $b = M^{p:r}$ we have,

$$\|b\|_\gamma = \|M\|^{p:r}_{p:s}, \quad \text{and} \quad \|a\|_{1/(1-\gamma)} = \|N\|^{p:s}_{r:s}.$$

Also,

$$\|b\|_1 = \|M\|^{p:r}_{p:r} = A^{p:r} \quad \text{and} \quad \|a\|_\infty = \|N\|^{p:s}_\infty = B^{p:s}.$$

The last two follow from Theorem 3.1 because $r < p$ and $q \leq s$. Using these four equations, Observations (2) and (3) give the upper bounds from the last line of the theorem. This completes the proof in the case $1 < q \leq s < r < p < \infty$. Corollary 4.4 shows that it remains valid for $0 < q \leq s < r < p \leq \infty$. □

A similar theorem covering the case $s < q < p \leq r$ follows by duality.

Theorem 5.9. *Suppose (3.1), (4.1), X and Y are purely atomic, C is the best constant in (5.1), and $s < q < p \leq r$. Then $C = GPP_{p:s/q:s}(M^{p:s}, N^{q:s})^{1/p:s}$. Consequently,*

$$\|M^*N^*\|_{p:s} \leq C \leq \min(A\|N\|_{p:s}, B\|M\|_{p:q}).$$

The last remaining case of the inclusion problem for purely atomic measures is the index range $s < q < r < p$. Some necessary and some sufficient conditions for inclusion in this case, already established in Theorems 4.8, 4.9 and 5.3, are stated in Theorem 4.1.

6. Multivariable Mixed Norm Lebesgue Space Inclusions

Let $p_1, \dots, p_n; r_1, \dots, r_n \in [1, \infty]$, $P = (L^{p_1}_{\mu_1}, \dots, L^{p_n}_{\mu_n})$ and $R = (L^{r_1}_{\kappa_1}, \dots, L^{r_n}_{\kappa_n})$ be the mixed norm spaces introduced in Section 2 for which the norms in each variable are Lebesgue norms. Let σ be a permutation of $\{1, \dots, n\}$ and let C_m and C be defined as in (2.1) and (2.2).

We restrict the Lebesgue indices to be at least 1 in this section to ensure that the Lebegue norms are Banach function norms. Using a result analogous to Lemma 4.2 it is easy to show that all results remain valid for any positive indices.

As in the two-variable case, for a large range of indices the Minkowski integral inequality together with single-variable inclusions will suffice to prove the mixed norm inclusion.

Theorem 6.1. *Let P, R and σ be as above. Suppose $P_i \subseteq R_i$ for $i = 1, \dots n$ and $r_i \leq p_j$ whenever $i < j$ and $\sigma^{-1}(i) > \sigma^{-1}(j)$. Then $\sigma(P) \subseteq R$ and $C = C_1 \dots C_n$.*

Proof. Fix $f \in L^+_\mu$. Let $R^{(k)}$ denote the $(n-k)$-vector obtained by removing $R_{\sigma(1)}, R_{\sigma(2)}, \dots, R_{\sigma(k)}$ from $R = (R_1, \dots, R_n)$, without disturbing the order of the remaining R's.

The single variable inclusion $P_{\sigma(1)} \subseteq R_{\sigma(1)}$ shows that

$$\|f\|_{(R_1,\dots,R_{\sigma(1)-1},R_{\sigma(1)},R_{\sigma(1)+1},\dots,R_n)} \leq C_{\sigma(1)}\|f\|_{(R_1,\dots,R_{\sigma(1)-1},P_{\sigma(1)},R_{\sigma(1)+1},\dots,R_n)}.$$

Next we move $P_{\sigma(1)}$ to the left by successively interchanging it with R_i for $i = \sigma(1) - 1$ down to $i = 1$. For each $i < \sigma(1)$ we have $\sigma^{-1}(i) > 1 = \sigma^{-1}(\sigma(1))$ so $r_i \leq p_{\sigma(1)}$ and thus, by Minkowski's integral inequality, these interchanges do not decrease the norm of f. Therefore,

$$\|f\|_R \leq C_{\sigma(1)}\|f\|_{(P_{\sigma(1)},R_1,\dots,R_{\sigma(1)-1},R_{\sigma(1)+1},\dots,R_n)} = C_{\sigma(1)}\|f\|_{(P_{\sigma(1)},R^{(1)})}.$$

To continue, locate $R_{\sigma(2)}$ in $R^{(1)}$, apply the one-variable inclusion $P_{\sigma(2)} \subseteq R_{\sigma(2)}$ and then interchange $P_{\sigma(2)}$ with each entry of $R^{(1)}$ to its left. For each such entry R_i, $i < \sigma(2)$ and $\sigma^{-1}(i) > 2 = \sigma^{-1}(\sigma(2))$, because $R_{\sigma(1)}$ is not in $R^{(1)}$. So $r_i \leq p_{\sigma(2)}$ and the interchanges do not decrease the norm. The result is,

$$\|f\|_R \leq C_{\sigma(1)}C_{\sigma(2)}\|f\|_{(P_{\sigma(1)},P_{\sigma(2)},R^{(2)})}.$$

Continuing in this way, we get,

$$\|f\|_R \le C_{\sigma(1)}C_{\sigma(2)}\dots C_{\sigma(n)}\|f\|_{(P_{\sigma(1)},P_{\sigma(2)},\dots,P_{\sigma(n)})} = C_1\dots C_n\|f\|_{\sigma(P)}.$$

Since, by hypothesis, each of $C_1,\dots,C_n$ is finite we have $\sigma(P)\subseteq R$ and $C\le C_1\dots C_n$. But Theorem 2.2 shows that $C_1\dots C_n\le C$ so we have $C=C_1\dots C_n$ to complete the proof. □

The restrictions on indices that enabled us to use the Minkowski integral inequality in the last proof are necessary conditions when there are no atomic measures involved. We can see this by applying our results in the two-variable case to the two-variable subinclusions from Lemma 2.3.

Theorem 6.2. *Let P, R and σ be as above. If none of the spaces $X_1,\dots,X_n$ is purely atomic, then $\sigma(P)\subseteq R$ if and only if all two-variable subinclusions hold, that is, if and only if*

$$(P_i,P_j)\subseteq(R_i,R_j),\quad \textit{when } i<j \textit{ and } \sigma^{-1}(i)<\sigma^{-1}(j), \textit{ and}$$
$$(P_j,P_i)\subseteq(R_i,R_j),\quad \textit{when } i<j \textit{ and } \sigma^{-1}(i)>\sigma^{-1}(j).$$

In this case, $C=C_1\dots C_n$.

Proof. The necessity of the two-variable subinclusions follows from Lemma 2.3. For the sufficiency, first observe that by Theorem 2.2, all single variable inclusions $P_i\subseteq R_i$ follow from the two variable subinclusions. Since the underlying spaces are not purely atomic, Theorem 3.1 shows that $r_i\le p_i$ for each i. Moreover, if $i<j$ and $\sigma^{-1}(i)>\sigma^{-1}(j)$, the subinclusion $(P_j,P_i)\subseteq(R_i,R_j)$ and Theorem 4.7 imply the index condition $\min(r_i,p_i)\le\max(r_j,p_j)$, which simplifies to $r_i\le p_j$. Thus, the conditions of Theorem 6.1 are satisfied. Its conclusions complete the proof. □

When we restrict our attention to mixed norm spaces based on (unweighted) ℓ^p-norms on infinite sequences we obtain a similar result—multivariable inclusions hold if and only if all two-variable subinclusions hold.

Theorem 6.3. *Let $X_1,\dots,X_n$ be countably infinite sets with counting measure, let $p_1,\dots,p_n;r_1,\dots,r_n\in[1,\infty]$ and let σ be a permutation of $\{1,\dots,n\}$. The mixed norm inclusion $(\ell^{p_{\sigma(1)}}(X_{\sigma(1)}),\dots,\ell^{p_{\sigma(n)}}(X_{\sigma(n)}))\subseteq(\ell^{r_1}(X_1),\dots,\ell^{r_n}(X_n))$ holds if and only if all two-variable subinclusions hold, that is, if and only if*

$$(\ell^{p_i}(X_i),\ell^{p_j}(X_j))\subseteq(\ell^{r_i}(X_i),\ell^{r_j}(X_j)),\quad \textit{when } i<j \textit{ and } \sigma^{-1}(i)<\sigma^{-1}(j), \textit{ and}$$
$$(\ell^{p_j}(X_j),\ell^{p_i}(X_i))\subseteq(\ell^{r_i}(X_i),\ell^{r_j}(X_j)),\quad \textit{when } i<j \textit{ and } \sigma^{-1}(i)>\sigma^{-1}(j).$$

In this case, $C=1$.

Proof. The necessity of the two-variable subinclusions follows from Lemma 2.3. For the sufficiency, first observe that by Theorem 2.2, all single variable inclusions $\ell^{p_i}(X_i)\subseteq\ell^{r_i}(X_i)$ follow from the two variable subinclusions. Theorem 3.1 shows that $r_i\ge p_i$ for each i and that $C_i=1$.

Suppose that $i<j$ and $\sigma^{-1}(i)>\sigma^{-1}(j)$. Then we have $(\ell^{p_j}(X_j),\ell^{p_i}(X_i))\subseteq(\ell^{r_i}(X_i),\ell^{r_j}(X_j))$ and Corollary 5.4 yields $p_i\le r_j$.

Fix a non-negative f. Let $P_m=\ell^{p_m}(X_m)$ and $R_m=\ell^{r_m}(X_m)$. This time, let $R^{(k)}$ be the $(n-k)$-vector obtained by removing $R_{\sigma(k)},R_{\sigma(k+1)},\dots,R_{\sigma(n)}$ from $R=(R_1,\dots,R_n)$, without disturbing the order of the remaining R's.

The constant is 1 in the single variable inclusion $P_{\sigma(n)}\subseteq R_{\sigma(n)}$, so

$$\|f\|_{(R_1,\dots,R_{\sigma(n)-1},R_{\sigma(n)},R_{\sigma(n)+1},\dots,R_n)}\le\|f\|_{(R_1,\dots,R_{\sigma(n)-1},P_{\sigma(n)},R_{\sigma(n)+1},\dots,R_n)}.$$

Next we move $P_{\sigma(n)}$ to the right by successively interchanging it with R_j for $j = \sigma(1)+1$ up to $j = n$. For each $j > \sigma(n)$ we have $\sigma^{-1}(j) < n = \sigma^{-1}(\sigma(n))$ so $p_{\sigma(n)} \le r_j$ and thus, by Minkowski's integral inequality (applied to sums), these interchanges do not decrease the norm of f. Therefore,

$$\|f\|_R \le \|f\|_{(R_1,\ldots,R_{\sigma(n)-1},R_{\sigma(n)+1},\ldots,R_n,P_{\sigma(n)})} = \|f\|_{(R^{(n)},P_{\sigma(n)})}.$$

To continue, locate $R_{\sigma(n-1)}$ in $R^{(n)}$, apply the one-variable inclusion $P_{\sigma(n-1)} \subseteq R_{\sigma(n-1)}$ and then interchange $P_{\sigma(n-1)}$ with each entry of $R^{(n)}$ to its right. For each such entry R_j, $j > \sigma(n-1)$ and $\sigma^{-1}(j) < n-1 = \sigma^{-1}(\sigma(n-1))$, because $R_{\sigma(n)}$ is not in $R^{(n)}$. So, again, we have $p_{\sigma(n-1)} \le r_j$ and the interchanges do not decrease the norm. The result is,

$$\|f\|_R \le \|f\|_{(R^{(n-1)},P_{\sigma(n-1)},P_{\sigma(n)})}.$$

Continuing in this way, we get,

$$\|f\|_R \le \|f\|_{(P_{\sigma(1)},P_{\sigma(2)},\ldots,P_{\sigma(n)})}.$$

That is,

$$\|f\|_{(\ell^{r_1}(X_1),\ldots,\ell^{r_n}(X_n))} \le \|f\|_{(\ell^{p_{\sigma(1)}}(X_{\sigma(1)}),\ldots,\ell^{p_{\sigma(n)}}(X_{\sigma(n)}))}.$$

Thus the inclusion, $(\ell^{p_{\sigma(1)}}(X_{\sigma(1)}),\ldots,\ell^{p_{\sigma(n)}}(X_{\sigma(n)})) \subseteq (\ell^{r_1}(X_1),\ldots,\ell^{r_n}(X_n))$, holds and $C \le 1$. Taking f to be supported on a single point shows that $C \ge 1$ so $C = 1$. $\square$

References

[1] R. Algervik and V. I. Kolyada, On Fournier-Gagliardo mixed norm spaces, Ann. Acad. Sci. Fenn. Math. **36**(2011), 493–508.

[2] S. Barza, A. Kamińska, L. E. Persson, and J. Soria, Mixed norm and multidimensional Lorentz spaces, Positivity **10**(2006), 539–554.

[3] A. Benedek and R. Panzone, The spaces L^p with mixed norm, Duke Math. J. **26**(1961), 301–324.

[4] C. Bennett and R. Sharpley, *Interpolation of Operators,* Academic Press, Orlando, Florida, U.S.A., 1988.

[5] A. P. Blozinski, Multivariate rearrangements and Banach function spaces with mixed norms, Trans. Amer. Math. Soc. **263**(1981), 149–167.

[6] A. Boccuto, A. V. Bukhvalov, and A. R. Sambucini, Some Inequalities in Classical Spaces with Mixed Norms, Positivity **6**(2002), 393–411.

[7] V. I. Bogachev, *Measure Theory*, Springer-Verlag, Berlin, Heidelberg, 2007.

[8] N. Clavero and J. Soria Mixed norm spaces and rearrangement invariant estimates J. Math. Anal. Appl., **419**(2014), 878–903.

[9] A. Defant, D. Popa and U. Schwarting, Coordinatewise multiple summing operators in Banach spaces, J. Funct. Anal., **259**(2010), 220–242.

[10] A. Defant, L. Frerick, J. Ortega-Cerdà, M. Ounaïes, and K. Seip, The Bohnenblust-Hille inequality for homogeneous polynomials is hypercontractive, Ann. of Math. (2) **174**(2011), 485–497.

[11] J. J. F. Fournier, Mixed norms and rearrangements: Sobolev's inequality and Littlewood's inequality, Ann. Mat. Pur. Appl., **148**(1987), 51–76.

[12] V. I. Kolyada, Iterated rearrangements and Gagliardo-Sobolev type inequalities, J. Math. Anal. Appl., **387**(2012), 335–348.

[13] W. A. J. Luxemburg, On the Measurability of a Function which Occurs in a Paper by A. C. Zaanen, Proc. Netherl. Acad. Sci. (A), **61**(1958), 259–265. (Indag. Math. **20**(1958), 259–265).

[14] W. A. J. Luxemburg, Addendum to On the Measurability of a Function which Occurs in a Paper by A. C. Zaanen, Proc. Netherl. Acad. Sci. (A), **66**(1963), 587–590. (Indag. Math. **25**(1963), 587–590).

[15] M. Milman, Notes on interpolation of mixed norm spaces and applications, Quart. J. Math. Oxford (2), **42**(1991), 325–334.

[16] D. Popa and G. Sinnamon, Blei's inequality and coordinatewise multiple summing operators, Publ. Mat., **57**(2013), 455–475.

[17] U. Schwarting, Vector Valued Bohnenblust-Hille Inequalities, Doctoral Thesis, Carl von Ossietzky University, Oldenburg, Germany, 2013.

[18] M. Väth, Some Measurability Results and Applications to Spaces with Mixed Family-norm, Positivity, **10**(2006), 737–753.

[19] A. C. Zaanen, *Integration*, North-Holland, Amsterdam, 1967.

Department of Mathematics, University of Western Ontario, London, Canada
E-mail address: wgrey@uwo.ca

Department of Mathematics, University of Western Ontario, London, Canada
E-mail address: sinnamon@uwo.ca