

## DISTANCES BETWEEN TWO-STATE MARKOV PROCESSES ATTAINABLE BY MARKOV JOININGS<sup>1</sup>

BY

MARTIN H. ELLIS

**ABSTRACT.** The function which assigns to each pair of two-state Markov processes the set of partition distances between them attainable by a Markov process on their joint atoms is computed. It is found that the infimum of these distances, the "Markov distance" between the pair, fails to satisfy the Triangle Inequality, hence fails to be a metric; thus in some cases the  $d$ -distance between two two-state Markov processes cannot be attained by a Markov process on their joint atoms.

**Introduction.** A stationary random process can be thought of as a 1-1 bimeasurable measure-preserving transformation on a probability space together with a finite measurable partition of the space (see [8] or [9] for a more detailed description). By "process" we will always mean stationary random process.

A process is a  $k$ -state process if its partition consists of  $k$  sets. Henceforth we will call sets in the partition "states" or "atoms", we assume that the states are indexed by the set  $\{1, 2, \dots, k\}$  unless otherwise indicated. A process is an  $n$ -step Markov process if for any two sequences of states  $S_1$  and  $S_2$  of length  $\geq n$  which have positive probability and which agree on their last  $n$  states, the probability of entering any state upon completing  $S_1$  equals the probability of entering that state upon completing  $S_2$ . A zero-step Markov process is called a Bernoulli process, and a one-step Markov process is simply called a Markov process.

The  $\bar{d}$ -distance, a natural and useful measure of the difference between two processes with the same collection of states, can be defined as follows: Given two processes  $\mathcal{T}_1$  and  $\mathcal{T}_2$  each with states  $\{1, 2, \dots, k\}$  let  $Y(\mathcal{T}_1, \mathcal{T}_2)$  be the set of all processes  $\mathcal{T}$  with states  $\{1, 2, \dots, k\} \times \{1, 2, \dots, k\}$  which have  $\mathcal{T}_1$  and  $\mathcal{T}_2$  as marginals. For each  $\mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2)$  let  $d_{\mathcal{T}}$  be the measure  $\mathcal{T}$

---

Received by the editors April 22, 1976 and, in revised form, January 13, 1977.

*AMS (MOS) subject classifications* (1970). Primary 28A65, 60J10; Secondary 60G35.

**Key words and phrases.** Stationary stochastic process, Markov process, joint process, partition distance,  $\bar{d}$ -distance, Markov joining, attaining  $\bar{d}$ , Markov joining attaining  $\bar{d}$ , distances attainable by Markov joinings.

<sup>1</sup>This paper is the main part of the author's Ph.D. thesis prepared at Stanford University under the direction of Donald S. Ornstein.

© American Mathematical Society 1978

assigns to the set of points in different  $\mathcal{T}_1$  and  $\mathcal{T}_2$  states (i.e.,  $d_{\mathcal{T}}$  is the sum of the probabilities  $\mathcal{T}$  assigns to its "off-diagonal" states), and let  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2) = \{d_{\mathcal{T}}: \mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2)\}$ ;  $d_{\mathcal{T}}$  is the distance between  $\mathcal{T}_1$  and  $\mathcal{T}_2$  attained by  $\mathcal{T}$ . The  $\bar{d}$ -distance between  $\mathcal{T}_1$  and  $\mathcal{T}_2$  is the infimum of  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2)$ ; it equals the minimum of  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2)$  since  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2)$  is a closed set.

In other words, the  $\bar{d}$ -distance between two processes is the smallest partition distance attainable between their partitions when embedding both processes in a third process; we can assume that the partition associated with the third process is the join of the partitions associated with the two embedded processes. (See Ornstein [8] or [9] for further discussion of  $\bar{d}$ .)

Given two processes  $\mathcal{T}_1$  and  $\mathcal{T}_2$  with the same number of states one can ask:

What is the  $\bar{d}$ -distance between them?

By what kinds of processes in  $Y(\mathcal{T}_1, \mathcal{T}_2)$  can this distance be attained?

What distances in  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2)$  can be attained by processes in  $Y(\mathcal{T}_1, \mathcal{T}_2)$  of a particular kind?

When  $\mathcal{T}_1$  and  $\mathcal{T}_2$  are Bernoulli processes, every distance in  $\mathcal{D}(\mathcal{T}_1, \mathcal{T}_2)$  is attained by a Bernoulli process in  $Y(\mathcal{T}_1, \mathcal{T}_2)$ . This can be seen as follows. If  $\mathcal{T}_1$  and  $\mathcal{T}_2$  are Bernoulli with states  $\{1, \dots, k\}$  and  $\mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2)$ , let  $\mathcal{T}^*$  be the Bernoulli process with states  $\{1, \dots, k\} \times \{1, \dots, k\}$  which assigns these joint states the same measure as  $\mathcal{T}$ ; then  $\mathcal{T}^* \in Y(\mathcal{T}_1, \mathcal{T}_2)$  and  $d_{\mathcal{T}^*} = d_{\mathcal{T}}$ . Furthermore if  $\mathcal{T}_1$  and  $\mathcal{T}_2$  assign to their states measures  $a_1, \dots, a_k$  and  $b_1, \dots, b_k$  respectively, then it is not hard to show that

$$D(\mathcal{T}_1, \mathcal{T}_2) = \left[ 1 - \sum_{i=1}^k \min\{a_i, b_i\}, \min\{1, 2 - \max\{a_i + b_i: i = 1, \dots, k\}\} \right].$$

In particular,  $\bar{d}(\mathcal{T}_1, \mathcal{T}_2) = 1 - \sum_{i=1}^k \min\{a_i, b_i\}$ , and this distance is attained by a Bernoulli process on the joint atoms.

In the paper we shall examine what distances are attainable between Markov processes by Markov processes on their joint atoms.

**DEFINITION.** Given two Markov processes  $\mathcal{T}_1$  and  $\mathcal{T}_2$  let  $\mathcal{M}(\mathcal{T}_1, \mathcal{T}_2) = \{d_{\mathcal{T}}: \mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2) \text{ and } \mathcal{T} \text{ is a Markov process}\}$  and let  $M(\mathcal{T}_1, \mathcal{T}_2) = \inf \mathcal{M}(\mathcal{T}_1, \mathcal{T}_2)$ .

$M(\mathcal{T}_1, \mathcal{T}_2)$  is the "Markov distance" between  $\mathcal{T}_1$  and  $\mathcal{T}_2$ .  $\mathcal{M}(\mathcal{T}_1, \mathcal{T}_2)$  is a closed set, so  $M(\mathcal{T}_1, \mathcal{T}_2) = d_{\mathcal{T}}$  for some Markov  $\mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2)$ .

Our aim in this paper is to compute  $\mathcal{M}$  and  $M$  for all pairs of two-state Markov processes. We assume each process has a first and a second state.

Actually we shall only compute  $\mathcal{M}$  and  $M$  for pairs where both processes have nonzero entropy—for the remaining pairs  $\mathcal{M}$  and  $M$  are not hard to compute and we have  $\mathcal{M} = \mathcal{D}$ , whence  $M = \bar{d}$ .<sup>2</sup> Throughout this paper we

will assume that all two-state Markov processes under consideration have positive entropy. Any such process is specified by its transition matrix

$$\begin{pmatrix} 1 - \kappa & \kappa \\ \lambda & 1 - \lambda \end{pmatrix}$$

where  $\kappa$  is the probability of leaving the first state and going to the second state,  $\lambda$  the probability of leaving the second state and going to the first. The requirement that the process have positive entropy reduces to having  $0 < \min(\kappa, \lambda) < 1$ ; we then have the measure of the first state equal to  $\lambda/(\kappa + \lambda)$ , the measure of the second state equal to  $\kappa/(\kappa + \lambda)$ .

Henceforth we will denote a two-state Markov process by its pair  $(\kappa, \lambda)$ , and we will always assume that  $0 < \min(\kappa, \lambda) < 1$ . We will think of  $\mathfrak{N}$  and  $M$  as functions on the collection of pairs  $((\alpha, \beta), (\gamma, \delta))$ ; to avoid confusion when considering a pair  $((\alpha, \beta), (\gamma, \delta))$  we will denote the first state of  $(\alpha, \beta)$  by  $A$ , the second state of  $(\alpha, \beta)$  by  $B$  and the first state of  $(\gamma, \delta)$  by  $C$ , the second state of  $(\gamma, \delta)$  by  $D$ . The states of  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  will be 1, 2, 3, 4, where  $1 = A \cap C$ ,  $2 = A \cap D$ ,  $3 = B \cap C$  and  $4 = B \cap D$ ;  $\mu$  will denote the measure. Thus  $d_{\mathfrak{T}} = \mu_{\mathfrak{T}}(2 \cup 3)$ , so attaining  $\bar{d}((\alpha, \beta), (\gamma, \delta))$  corresponds to minimizing  $\mu_{\mathfrak{T}}(2 \cup 3)$ ,  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$ .

We will find that  $M$  sometimes fails to satisfy the Triangle Inequality and hence is not a metric, whence  $M \neq \bar{d}$ . We will also see that  $\mathfrak{N}((\alpha, \beta), (\gamma, \delta))$  need not form an interval. In the appendix we state some results concerning non-Markovian joinings.

**1. Definitions and lemmas.** In §1 we derive the Basic Lemma, our criterion for a four-state Markov process to yield a two-state Markov process when its atoms are lumped into two states consisting of two atoms each. For each pair of two-state Markov processes  $((\alpha, \beta), (\gamma, \delta))$  we introduce the simple processes for  $((\alpha, \beta), (\gamma, \delta))$ , a collection of Markov processes in  $Y((\alpha, \beta), (\gamma, \delta))$ , and find the set of partition distances between  $(\alpha, \beta)$  and  $(\gamma, \delta)$  attained by simple processes for  $((\alpha, \beta), (\gamma, \delta))$ . This section ends by using symmetry to make assumptions which will simplify the computation of  $\mathfrak{N}$  and  $M$ . Preliminaries completed,  $\mathfrak{N}$  and  $M$  are computed in §§2 and 3.

If  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  is Markov, then  $\mathfrak{T}$  is a four-state Markov process with partition  $\{1, 2, 3, 4\}$  such that when atoms 1 and 2 are lumped together and atoms 3 and 4 are lumped together (respectively, atoms 1 and 3 and atoms 2 and 4), the resulting two-state process is isomorphic to the Markov

---

<sup>2</sup>If one process has positive entropy and the other zero entropy then  $\mathfrak{N}$  and  $\mathfrak{D}$  both consist of the single value which is the distance attained by independent matching (this can be shown directly or by observing that the positive entropy process is a  $K$ -automorphism and hence in any joint embedding of the two processes their  $\sigma$ -algebras, and in particular their partitions, must be orthogonal (see [1])); if both have entropy zero the computation of  $\mathfrak{N}$  and  $M$  is trivial.

process  $(\alpha, \beta)$  (respectively,  $(\gamma, \delta)$ ). We begin our analysis by establishing necessary and sufficient conditions (Lemma 1.3) which a Markov process with partition  $\{k_1, k_2, l_1, l_2\}$  must satisfy for the lumped process with partition  $\{K, L\}$ ,  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$  to be a Markov process.

**DEFINITION.** Given a four-state Markov process  $\mathfrak{T}$  with transformation  $T$ , measure  $\mu$  and partition  $P = \{k_1, k_2, l_1, l_2\}$ , if  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$ :

$K$  is *level* if  $\mu(T(k_1) \cap L): \mu(k_1) = \mu(T(k_2) \cap L): \mu(k_2)$ ,

$K$  is *respectful* if  $\mu(T(K) \cap l_1): \mu(l_1) = \mu(T(K) \cap l_2): \mu(l_2)$ ,

$K$  is *uniform* if for  $i \in \{1, 2\}$ ,  $\mu(T(k_i) \cap l_1): \mu(l_1) = \mu(T(k_i) \cap l_2): \mu(l_2)$  (uniform is a strengthening of respectful),

$K$  is *perfect* if it is both level and uniform,

$K$  is *balanced* if  $\mu(T(K) \cap k_1): \mu(k_1) = \mu(T(K) \cap k_2): \mu(k_2)$ .

(Symmetric definitions apply for  $L$ :  $L$  is level if  $\mu(T(l_1) \cap K): \mu(l_1) = \mu(T(l_2) \cap K): \mu(l_2)$ , etc.)

Whenever either of the ratios is 0:0, the equation will be said to hold.

In words,  $K$  is level if each of  $K$ 's atoms sends the same proportion of its measure to  $L$ ,  $K$  is respectful if measure entering  $L$  from  $K$  distributes itself between the atoms of  $L$  in proportion to their measure (that is,  $K$  respects  $L$ 's distribution), and  $K$  is uniform if each atom of  $K$  respects  $L$ 's distribution.  $K$  is balanced if  $K$  respects its own distribution (that is,  $K$  maintains the proper balance between its atoms). The above definitions imply that  $K$  is perfect if and only if for  $i, j \in \{1, 2\}$ ,  $\mu(T(k_i) \cap l_j) = (\kappa + \lambda) \mu(k_i) \mu(l_j)$ , where  $\kappa = \mu(T(K) \cap L)/\mu(K)$  and  $\lambda = \mu(T(L) \cap K)/\mu(L)$ .

**Notation.** For  $n$  and  $m$  measurable subsets of the measure space, let  $e_{nm} = \mu(T(n) \cap m)/\mu(n)$  if  $\mu(n) > 0$ ; if  $\mu(n) = 0$ ,  $e_{nm}$  is undefined (or can be thought of as any number we wish).

**EXAMPLES.** In each of these examples we give part of the transition matrix for a Markov process with partition  $\{k_1, k_2, l_1, l_2\}$ ,  $\mu(k_1) = .2$ ,  $\mu(k_2) = .4$ ,  $\mu(l_1) = .1$ ,  $\mu(l_2) = .3$ ,  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$ ,  $e_{KL} = .6$ , and  $e_{LK} = .9$  (the states are ordered  $k_1, k_2, l_1, l_2$ ).

$$\begin{bmatrix} .3 & .2 & .1 & .4 \\ .1 & .1 & .3 & .5 \\ & & \vdots & \end{bmatrix} \quad K \text{ is } \begin{array}{l} \text{level} \\ \text{not respectful} \\ \text{not balanced,} \end{array}$$

$$\begin{bmatrix} .7 & 0 & .25 & .05 \\ .15 & .1 & .1 & .65 \\ & & \vdots & \end{bmatrix} \quad K \text{ is } \begin{array}{l} \text{not level} \\ \text{respectful} \\ \text{not uniform} \\ \text{not balanced,} \end{array}$$

|   |        |   |
|---|--------|---|
| $\begin{bmatrix} .4 & .4 & .05 & .15 \\ .1 & .1 & .20 & .60 \\ & & \vdots & \\ & & & \end{bmatrix}$ | $K$ is | not level<br>uniform<br>not balanced,               |
| $\begin{bmatrix} .2 & .2 & .15 & .45 \\ .2 & .2 & .15 & .45 \\ & & \vdots & \\ & & & \end{bmatrix}$ | $K$ is | perfect<br>not balanced,                            |
| $\begin{bmatrix} .2 & .2 & .15 & .45 \\ .1 & .3 & .15 & .45 \\ & & \vdots & \\ & & & \end{bmatrix}$ | $K$ is | perfect<br>balanced,                                |
| $\begin{bmatrix} 0 & .6 & .25 & .15 \\ .2 & .1 & .1 & .6 \\ & & \vdots & \\ & & & \end{bmatrix}$    | $K$ is | not level<br>respectful<br>not uniform<br>balanced. |

In general, all of the twelve combinations of level, respectful, uniform, and balanced that are not by their definitions immediately eliminated are possible (however, for example, if  $k_1$  or  $k_2$  has measure zero,  $K$  will automatically be level and balanced, and if  $K$  is respectful,  $K$  will automatically be uniform).

Since the transformations under consideration are bijections we have

LEMMA 1.1.  *$K$  is respectful iff  $L$  is balanced.*

LEMMA 1.2. *Given a Markov process with transformation  $T$  and partition  $\{k_1, k_2, l_1, l_2\}$ ,  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$ , if the  $\{K, L\}$ -process is Markov and  $K$  is not level, then  $K$  is balanced.*

PROOF. Suppose  $\mu(T(K) \cap L) = \kappa\mu(K)$ . If  $K$  is not balanced then  $TK \cap K$  is skewed in  $K$  (i.e.  $\mu(TK \cap k_1) : \mu(k_1) \neq \mu(TK \cap k_2) : \mu(k_2)$ ), and if  $K$  is not level, any skewed distribution in  $K$  does not send  $\kappa$  of its measure to  $L$  (it sends more or less than  $\kappa$  depending on whether the skewing in  $K$  favors the atom sending more or less than  $\kappa$  of its measure to  $L$ ). Q.E.D.

The following lemma will be the basic tool for our analysis, for it indicates what conditions a four-state Markov process must satisfy to be in  $Y((\alpha, \beta), (\gamma, \delta))$ .

LEMMA 1.3 (BASIC LEMMA). *Given Markov process  $\mathcal{T}$  with transformation  $T$  and partition  $\{k_1, k_2, l_1, l_2\}$ ,  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$ , the  $\{K, L\}$ -process is Markov if and only if at least one of the following conditions holds:*

- (1) *Both  $K$  and  $L$  are level.*
- (2) *Both  $K$  and  $L$  are respectful and balanced.*

(3)  $K$  is perfect and  $L$  balanced, or  $L$  is perfect and  $K$  is balanced.

PROOF. The  $\{K, L\}$ -process is Markov if and only if for every positive integer  $m$  and for every sequence  $P_1, \dots, P_m$  of  $K$ 's and  $L$ 's, if  $S = T^1 P_1 \cap T^2 P_2 \cap \dots \cap T^m P_m$ ,  $E = K \cap S$ ,  $F = L \cap S$ ,

$$e_{EL} = e_{KL} \quad (1.1)$$

and

$$e_{FK} = e_{LK}. \quad (1.2)$$

(Whenever  $e_{EL}$  or  $e_{FK}$  is undefined, the equation will be said to hold.) The behavior of  $E$  under  $T$  depends only upon the distribution of  $K \cap S$  between  $k_1$  and  $k_2$ . If  $K$  is level, no matter what distribution  $E$  has, the same proportion of  $E$  will be sent to  $L$  under  $T$  as  $K$  sends to  $L$  under  $T$ , so for all histories  $S$  (1.1) will be satisfied. If  $K$  is not level, however, then (1.1) will fail to hold for  $S$  unless  $\mu(k_1 \cap S) : \mu(k_2 \cap S) = \mu(k_1) : \mu(k_2)$ . Hence (1.1) is satisfied for all histories  $S$  iff

(1.3)  $K$  is level, or for every history  $S$ ,  $\mu(k_1 \cap S) : \mu(k_2 \cap S) = \mu(k_1) : \mu(k_2)$ .

Similarly, (1.2) is satisfied iff

(1.4)  $L$  is level, or for every history  $S$ ,  $\mu(l_1 \cap S) : \mu(l_2 \cap S) = \mu(l_1) : \mu(l_2)$ .

If (1) is satisfied, then (1.3) and (1.4) hold, so the  $\{K, L\}$ -process is Markov.

If (2) is satisfied, a straightforward induction on the length of histories shows that for every history  $S$ ,  $\mu(k_1 \cap S) : \mu(k_2 \cap S) = \mu(k_1) : \mu(k_2)$  and  $\mu(l_1 \cap S) : \mu(l_2 \cap S) = \mu(l_1) : \mu(l_2)$ , whence (1.3) and (1.4) hold, so the  $\{K, L\}$ -process is Markov.

If (3) is satisfied, say  $K$  is perfect and  $L$  balanced, then  $K$  is level so (1.3) holds; furthermore, for every history  $S$ ,  $\mu(l_1 \cap S) : \mu(l_2 \cap S) = \mu(l_1) : \mu(l_2)$ , for the uniformity of  $K$  insures that any distribution of measure in  $K$  enters  $L$  with the proper distribution between  $l_1$  and  $l_2$ , and since  $L$  is balanced it will maintain the proper distribution between  $l_1$  and  $l_2$ .

To show at least one of (1), (2), (3) is necessary, suppose neither (1) nor (2) hold. Since (1) does not hold, either  $K$  or  $L$  must not be level—say  $L$  is not level. Then by Lemma 1.2,  $L$  must be balanced, so by Lemma 1.1,  $K$  must be respectful. Now if  $K$  is balanced, by Lemma 1.1,  $L$  would be respectful, but then (2) would hold; thus  $K$  must not be balanced, whence by Lemma 1.2 we must have  $K$  level. Furthermore, since  $K$  is not balanced  $K \cap TK$  is skewed in  $K$ ; yet, since  $L$  is not level, by (1.4),  $\mu(l_1 \cap T(K \cap TK)) : \mu(l_2 \cap T(K \cap TK)) = \mu(l_1) : \mu(l_2)$ . Hence not only  $K$  enters  $L$  respectfully, but a skewed weighting of  $K$  enters  $L$  respectfully. Thus  $K$  must be uniform, so  $K$  is level and uniform, i.e., perfect and (3) holds. Q.E.D.

COMMENT. By Lemma 1.1 we can replace condition (2) with any of

(2')  $K$  and  $L$  are both respectful,

(2'')  $K$  and  $L$  are both balanced,

(2''')  $K$  is balanced and respectful or  $L$  is balanced and respectful,

and we can replace condition (3) by

(3')  $K$  is perfect or  $L$  is perfect.

LEMMA 1.4. *Given a Markov process with transformation  $T$  and partition  $\{k_1, k_2, l_1, l_2\}$ , if  $K = k_1 \cup k_2$ ,  $L = l_1 \cup l_2$  and the  $\{K, L\}$ -process is isomorphic to  $(\kappa, \lambda)$ , then for  $i \in \{1, 2\}$ ,*

$$\mu(k_i) \neq 0 \Rightarrow e_{k_i k_i} \leq 1 - \kappa, \quad \mu(l_i) \neq 0 \Rightarrow e_{l_i l_i} \leq 1 - \lambda.$$

PROOF.  $\forall n \in \mathbb{N}$

$$\begin{aligned} \mu(K \cap TK \cap \cdots \cap T^n K) &= \mu(K)(1 - \kappa)^n \\ &\geq \mu(k_i \cap Tk_i \cap \cdots \cap T^n k_i) = \mu(k_i)e_{k_i k_i}^n, \\ \mu(L \cap TL \cap \cdots \cap T^n L) &= \mu(L)(1 - \lambda)^n \\ &\geq \mu(l_i \cap Tl_i \cap \cdots \cap T^n l_i) = \mu(l_i)e_{l_i l_i}^n. \quad \text{Q.E.D.} \end{aligned}$$

DEFINITION. We will call  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  a *simple process* (for  $((\alpha, \beta), (\gamma, \delta))$ ) if  $\mathfrak{T}$  is Markov and  $A, B, C$ , and  $D$  are all level.

The *standard process* for  $((\alpha, \beta), (\gamma, \delta))$  is the simple process for  $((\alpha, \beta), (\gamma, \delta))$  which minimizes  $e_{i2} + e_{i3}$  for  $i \in \{1, 2, 3, 4\}$ . In other words, the standard process for  $((\alpha, \beta), (\gamma, \delta))$  is the simple process for  $((\alpha, \beta), (\gamma, \delta))$  which is "locally best possible" in the sense that each of its four states sends as small a proportion of its measure to  $2 \cup 3$  (the "bad states") as it can consistent with the process being a simple process for  $((\alpha, \beta), (\gamma, \delta))$ . It is not hard to show that the standard process for  $((\alpha, \beta), (\gamma, \delta))$  is unique and has transition matrix

$$\begin{bmatrix} 1 - \max(\alpha, \gamma) & \gamma - \min(\alpha, \gamma) & \alpha - \min(\alpha, \gamma) & \min(\alpha, \gamma) \\ \delta - \max(0, \alpha + \delta - 1) & \max(0, 1 - (\alpha + \delta)) & \max(0, \alpha + \delta - 1) & \alpha - \max(0, \alpha + \delta - 1) \\ \beta - \max(0, \beta + \gamma - 1) & \max(0, \beta + \gamma - 1) & \max(0, 1 - (\beta + \gamma)) & \gamma - \max(0, \beta + \gamma - 1) \\ \min(\beta, \delta) & \beta - \min(\beta, \delta) & \delta - \min(\beta, \delta) & 1 - \max(\beta, \delta) \end{bmatrix}.$$

*Notation.* Let  $\text{St}((\alpha, \beta), (\gamma, \delta))$  denote  $\mu(2 \cup 3)$  for the standard process for  $((\alpha, \beta), (\gamma, \delta))$ .

LEMMA 1.5. (i)  $\text{St}((\alpha, \beta), (\gamma, \delta))$  is the minimum of  $\mu(2 \cup 3)$  for the simple processes for  $((\alpha, \beta), (\gamma, \delta))$ .

(ii)  $1 - \text{St}((\alpha, \beta), (\delta, \gamma))$  is the maximum of  $\mu(2 \cup 3)$  for the simple processes for  $((\alpha, \beta), (\gamma, \delta))$ .

(iii)  $[\text{St}((\alpha, \beta), (\gamma, \delta)), 1 - \text{St}((\alpha, \beta), (\delta, \gamma))] \subseteq \mathfrak{N}((\alpha, \beta), (\gamma, \delta))$ .

PROOF. (i) Note that

$$\mu(3) = \mu(2) + \mu(C) - \mu(A). \quad (1.5)$$

Thus, minimizing  $\mu(2 \cup 3)$  is equivalent to minimizing  $\mu(2)$  (which is equivalent to maximizing  $\mu(1)$  or maximizing  $\mu(4)$  or minimizing  $\mu(3)$ ). Assume  $\mu(C) - \mu(A) \geq 0$  (otherwise interchange  $(\alpha, \beta)$  and  $(\gamma, \delta)$ ).

Any simple process for  $((\alpha, \beta), (\gamma, \delta))$  with transition probabilities  $e_{ij}$  satisfies

$$\begin{aligned} \mu(T(1 \cup 4) \cap (2 \cup 3)) &= (e_{12} + e_{13})\mu(1) + (e_{42} + e_{43})\mu(4) \\ &= (\alpha + \gamma - 2e_{14})\mu(1) + (\beta + \delta - 2e_{41})\mu(4) \end{aligned}$$

and

$$\begin{aligned} \mu(T(2 \cup 3) \cap (1 \cup 4)) &= (e_{21} + e_{24})\mu(2) + (e_{31} + e_{34})\mu(3) \\ &= (\alpha + \delta - 2e_{23})\mu(2) + (\beta + \gamma - 2e_{32})\mu(3) \\ &= (\alpha + \beta + \gamma + \delta - 2e_{23} - 2e_{32})\mu(2) + (\beta + \gamma - 2e_{32})(\mu(C) - \mu(A)). \end{aligned}$$

Thus any simple process for  $((\alpha, \beta), (\gamma, \delta))$  satisfies

$$\begin{aligned} &(\alpha + \gamma - 2e_{14})\mu(1) + (\beta + \delta - 2e_{41})\mu(4) \\ &\quad + (2e_{32} - \beta - \gamma)(\mu(C) - \mu(A)) \\ &= (\alpha + \beta + \gamma + \delta - 2e_{23} - 2e_{32})\mu(2). \end{aligned}$$

Among all simple processes for  $((\alpha, \beta), (\gamma, \delta))$  the standard process is the one which minimizes  $(\alpha + \gamma - 2e_{14})$ ,  $(\beta + \delta - 2e_{41})$ ,  $(2e_{32} - \beta - \gamma)$ , and maximizes  $(\alpha + \beta + \gamma + \delta - 2e_{23} - 2e_{32})$ . If a simple process could improve upon the standard process,  $\mu(1)$  and  $\mu(4)$  would increase and  $\mu(2)$  would decrease, whence (unless  $\alpha = \gamma$  and  $\beta = \delta$ , in which case an improvement on the standard process is impossible) by the preceding observation the right side of equation (1.6) would have to decrease and the left side increase (from their values for the standard process), whence (1.6) would no longer hold.

(ii) Any  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  can be viewed as an element of  $Y((\alpha, \beta), (\delta, \gamma))$  and vice versa by interchanging states 1 and 2 and states 3 and 4. Simple processes in  $Y((\alpha, \beta), (\gamma, \delta))$  become simple processes in  $Y((\alpha, \beta), (\delta, \gamma))$  and vice versa. By part (i),  $\text{St}((\alpha, \beta), (\delta, \gamma))$  is the minimum of  $\mu(2 \cup 3) = \mu((A \cap C) \cup (B \cap D))$  among all simple processes in  $Y((\alpha, \beta), (\delta, \gamma))$ ; hence, by the preceding observation,  $\text{St}((\alpha, \beta), (\delta, \gamma))$  is the minimum of  $\mu((A \cap C) \cup (B \cap D))$  among all simple processes in  $Y((\alpha, \beta), (\gamma, \delta))$ . Thus

$$\begin{aligned} \max \mu(2 \cup 3) &= 1 - \min \mu(1 \cup 4) \\ &= 1 - \min \mu((A \cap C) \cup (B \cap D)) \\ &= 1 - \text{St}((\alpha, \beta), (\delta, \gamma)). \end{aligned}$$

(iii) Let  $U$  be the transition matrix for the standard process for  $((\alpha, \beta), (\gamma, \delta))$  and let  $V$  be the transition matrix for the simple process for  $((\alpha, \beta), (\gamma, \delta))$  which makes  $\mu(2 \cup 3) = 1 - \text{St}((\alpha, \beta), (\delta, \gamma))$ . For each  $t \in [0, 1]$ , let

$$P_t = (1 - t)U + tV.$$

Then for all  $t \in [0, 1]$ ,  $P_t$  gives a simple process for  $((\alpha, \beta), (\gamma, \delta))$ , and from equation (1.6) we see that as  $t$  increases from 0 to 1,  $\mu(2 \cup 3)$  increases monotonically and continuously from  $\text{St}((\alpha, \beta), (\gamma, \delta))$  to  $1 - \text{St}((\alpha, \beta), (\delta, \gamma))$ . Q.E.D.

*Digression.* For Markov  $\mathcal{T} \in Y(\mathcal{T}_1, \mathcal{T}_2)$ ,  $\mathcal{T}_1$  and  $\mathcal{T}_2$  two  $n$ -state Markov processes,  $n \geq 3$ , there are natural extensions of “level” and “simple process” and “standard process” (for  $n = 3$  there will still be a unique “standard process”, though for  $n \geq 4$  there may be more than one “standard process”, for when  $n \geq 4$  “locally best possible” may not completely determine the process). However, even for  $n = 3$  it sometimes is not the case that the minimum distance attained by the “simple processes” is attained by the (any) “standard process”. This happens because it is no longer the case that increasing the measure of any good (diagonal) state will not cause the increase in measure of some bad (off-diagonal) state(s) at the expense of some other good state(s), and sometimes a local “improvement” causes a global deterioration.

*Symmetry.* Notice that  $\mathcal{N}$ ,  $M$ ,  $\mathcal{D}$ ,  $\bar{d}$ , and  $\text{St}$  give the same values for each of the pairs  $((\alpha, \beta), (\gamma, \delta))$ ,  $((\beta, \alpha), (\delta, \gamma))$ ,  $((\gamma, \delta), (\alpha, \beta))$ ,  $((\delta, \gamma), (\beta, \alpha))$ .

DEFINITION. We will call four such pairs *equivalents* of each other.

Using this symmetry, we can simplify our proofs by assuming that when we are given a pair of processes  $((\alpha, \beta), (\gamma, \delta))$ , it is in some standard form.

DEFINITION.  $((\alpha, \beta), (\gamma, \delta))$  is in *proper form* (abbreviations: PF, PFI, PFII) if it is in one of the following two forms:

(I)  $((\alpha, \beta), (\alpha - u, \beta + w))$  with  $\mu(C) \geq \mu(A)$ ,  $w \geq u \geq 0$ , and if  $w = u$  then  $\gamma < \beta$ .

(II)  $((\alpha, \beta), (\alpha + u, \beta + w))$  with  $\mu(C) \geq \mu(A)$  and  $u, w > 0$ .

LEMMA 1.6. *Given any pair of processes  $((\alpha, \beta), (\gamma, \delta))$ ,*

(i) *at least one of its four equivalents will be in PF,*

(ii) *if it has an equivalent in PFI, then it does not have an equivalent in PFII and vice versa.*

PROOF. Straightforward.

Henceforth we shall assume that (unless stated otherwise)  $((\alpha, \beta), (\gamma, \delta))$  is in PF. Thus  $\mu(C) \geq \mu(A)$ , so by (1.5) we have

$$\mu(3) \geq \mu(2). \quad (1.7)$$

Given a Markov process  $\mathcal{T}$  let  $\mathcal{T}^{-1}$  denote the process with the same measure space and the same partition as  $\mathcal{T}$  but with  $T$  replaced by  $T^{-1}$ .  $\mathcal{T}^{-1}$  is a Markov process since the inverse of a Markov process is Markov, and all the states have the same probability for  $\mathcal{T}^{-1}$  as for  $\mathcal{T}$ .

**LEMMA 1.7.** *If  $\mathcal{T}$  is a Markov process with states  $\{k_1, k_2, l_1, l_2\}$  and  $(\kappa, \lambda)$  is embedded in  $\mathcal{T}$  with  $K = k_1 \cup k_2, L = l_1 \cup l_2$ , then  $(\kappa, \lambda)$  is embedded in  $\mathcal{T}^{-1}$  with  $K = k_1 \cup k_2, L = l_1 \cup l_2$ .*

**PROOF.** Since  $T$  is measure preserving,  $T^{-1}$  sends  $\kappa$  of the measure of  $K$  to  $L$  and  $\lambda$  of the measure of  $L$  to  $K$ , and  $T^{-1}$  is Markov with respect to the partition  $\{K, L\}$  since the inverse of the transformation which is Markov with respect to a partition is a Markov process with respect to that partition. Q.E.D.

Note that if  $\mathcal{T}$  satisfies condition  $(m)$  of the Basic Lemma ( $m \in \{1, 2, 3\}$ ), then  $\mathcal{T}^{-1}$  satisfies condition  $((3 - m) \bmod 3)$  of the Basic Lemma.

**LEMMA 1.8.** *For each  $d \in \mathcal{D}((\alpha, \beta), (\gamma, \delta))$ , if there is a Markov  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathcal{T}} = d$ , then for each of the states  $A, B, C, D$  there is a Markov  $\mathcal{T}' \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathcal{T}'} = d$  and with that state level for  $\mathcal{T}'$ , and if that state is not perfect for  $\mathcal{T}'$ , then its complement is also level for  $\mathcal{T}'$ .*

**PROOF.** Suppose we want  $C$  and  $D$  level or  $C$  perfect. Given a Markov  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathcal{T}} = d$ , if  $C$  and  $D$  are not level for  $\mathcal{T}$  and  $C$  is not perfect for  $\mathcal{T}$ , then  $\mathcal{T}$  satisfies either condition (2) of the Basic Lemma, in which case  $C$  and  $D$  are level for  $\mathcal{T}^{-1}$ , or  $\mathcal{T}$  satisfies condition (3) of the Basic Lemma with  $D$  perfect, in which case  $C$  is perfect for  $\mathcal{T}^{-1}$ , and, by Lemma 1.7,  $\mathcal{T}^{-1}$  is a Markov process in  $Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathcal{T}^{-1}} = d_{\mathcal{T}} = d$ . Q.E.D.

**2. Proper form I.** In §2 we determine  $\mathfrak{N}$  and  $M$  for all pairs in PFI. We find that for all pairs  $((\alpha, \beta), (\gamma, \delta))$  in PFI,  $M((\alpha, \beta), (\gamma, \delta)) = \text{St}((\alpha, \beta), (\gamma, \delta))$  unless  $\alpha + \beta = \gamma + \delta$ , in which case  $M((\alpha, \beta), (\gamma, \delta))$  is the partition distance. This shows us that  $M$  fails to satisfy the Triangle Inequality, hence is not a metric and is not identically  $\bar{d}$ .

The smallest distance we can try to attain by any process in  $Y((\alpha, \beta), (\gamma, \delta))$  is  $\mu(C) - \mu(A)$ , since  $\mu(3) = \mu(C) - \mu(A) + \mu(2)$ , and we can attain this iff we can attain  $\mu(2) = 0$ . Thus whenever we have attained  $\mu(2) = 0$  by a Markov  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$ , we know we have attained  $M((\alpha, \beta), (\gamma, \delta))$ , and furthermore that this equals  $\bar{d}((\alpha, \beta), (\gamma, \delta))$ .

**LEMMA 2.1.** *If  $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$ , then  $\alpha \geq \gamma$  and  $\delta \geq \beta$ .*

**PROOF.** If  $T$  is the transformation in which  $(\alpha, \beta)$  and  $(\gamma, \delta)$  are embedded with  $A \subseteq C$ , then

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \mu\left(\bigcap_{i=0}^n T^{-i}A\right) &= \mu(A)(1 - \alpha)^n \leq \mu\left(\bigcap_{i=0}^n T^{-i}C\right) \\ &= \mu(C)(1 - \gamma)^n, \end{aligned}$$

whence since  $\mu(A) > 0$ ,  $\alpha \geq \gamma$ , and

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \mu\left(\bigcap_{i=0}^n T^{-i}D\right) &= \mu(D)(1 - \delta)^n \leq \mu\left(\bigcap_{i=0}^n T^{-i}B\right) \\ &= \mu(B)(1 - \beta)^n, \end{aligned}$$

whence since  $\mu(D) > 0$ ,  $\delta \geq \beta$ . Q.E.D.

**COROLLARY 2.2.** *If  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$ , then  $\bar{d}((\alpha, \beta), (\gamma, \delta)) > \mu(C) - \mu(A)$ . (Hence  $M((\alpha, \beta), (\gamma, \delta)) > \mu(C) - \mu(A)$ .)*

For  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFI}$

$$\text{St}((\alpha, \beta), (\gamma, \delta)) = \begin{cases} \mu(C) - \mu(A) & \text{if } \beta + \gamma \leq 1, \\ (\mu(C) - \mu(A)) \frac{\beta + \gamma}{2 - (\beta + \gamma)} & \text{if } \beta + \gamma > 1. \end{cases} \quad (2.1)$$

Hence

**PROPOSITION 2.3.** *For  $((\alpha, \beta), (\gamma, \delta))$  in PFI, if  $\beta + \gamma \leq 1$ , then*

$$\begin{aligned} M((\alpha, \beta), (\gamma, \delta)) &= \text{St}((\alpha, \beta), (\gamma, \delta)) \\ &= \bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A). \end{aligned}$$

**LEMMA 2.4.** *If for a Markov  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$  either  $A$  or  $B$  is perfect and either  $C$  or  $D$  is perfect, then one of the following must hold:*

- (1) *the atom both perfect states contain has measure zero,*
- (2) *the atom neither perfect state contains has measure zero,*
- (3)  $\alpha + \beta = \gamma + \delta$ .

**PROOF.** Suppose for example  $A$  and  $C$  are perfect. If  $\mu(1) \neq 0$ , then

$$\begin{aligned} e_{14} &= (\alpha + \beta) \mu(4) && \text{since } A \text{ is perfect,} \\ &= (\gamma + \delta) \mu(4) && \text{since } C \text{ is perfect. Q.E.D.} \end{aligned}$$

**DEFINITION.**  $((\alpha, \beta), (\gamma, \delta))$  is in the *special case* if  $\alpha + \beta = \gamma + \delta$ .

**PROPOSITION 2.5.** *For  $((\alpha, \beta), (\gamma, \delta))$  in the special case,  $M((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$ . (Hence  $\bar{d}((\alpha, \beta), (\gamma, \delta)) = M((\alpha, \beta), (\gamma, \delta))$ .)*

**PROOF.** Since  $((\alpha, \beta), (\gamma, \delta))$  is in PF,  $\gamma \leq \min(\alpha, \beta)$  ( $\gamma < \alpha$  unless  $(\gamma, \delta) = (\alpha, \beta)$ ,  $\gamma < \beta$  unless  $(\gamma, \delta) = (\beta, \alpha)$ ). Then the Markov process with transition matrix

$$\begin{bmatrix} 1 - \alpha & 0 & \alpha - \gamma & \gamma \\ \vdots & \vdots & \vdots & \vdots \\ \beta - \gamma & 0 & 1 - \beta & \gamma \\ \delta & 0 & 0 & 1 - \delta \end{bmatrix}$$

attains this distance ( $e_{2i}$  need not be defined since  $\mu(2) = 0$ ). Note that all the entries in the matrix are nonnegative. For this process,  $A$  and  $C$  are perfect. Q.E.D.

*Digression.* For  $((\alpha, \beta), (\gamma, \delta))$  in the special case with  $(\alpha, \beta) \neq (\gamma, \delta)$  we have a one-parameter family of Markov processes in  $Y((\alpha, \beta), (\gamma, \delta))$  attaining  $A \subseteq C$  with  $C$  level: the family with  $A$  and  $C$  perfect and  $\mu(2) = 0$ . This family together with its family of inverses are the only Markov processes in  $Y((\alpha, \beta), (\gamma, \delta))$  attaining  $A \subseteq C$ . If  $\beta + \gamma > 1$ , these two collections are disjoint; if  $\beta + \gamma \leq 1$ , they have one member in common—the standard process.

The largest distance we can try to attain by any  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$  would occur if we could attain  $\mu(2) = \min\{\mu(A), \mu(D)\}$ . In the special case the minimum is  $\mu(D)$ ; if  $\mu(2) = \mu(D)$ , then  $\mu(2 \cup 3) = 2\mu(2) + \mu(C) - \mu(A) = 2\mu(D) + \mu(C) - \mu(A) = \mu(B) + \mu(D)$ . In the special case this distance can be attained by a Markov process in  $Y((\alpha, \beta), (\gamma, \delta))$ , and furthermore, every value between this and  $\mu(C) - \mu(A)$  is attained by a Markov process in  $Y((\alpha, \beta), (\gamma, \delta))$ :

**PROPOSITION 2.6.** *For  $((\alpha, \beta), (\gamma, \delta))$  in the special case,*

$$\mathcal{D}((\alpha, \beta), (\gamma, \delta)) = \mathcal{N}((\alpha, \beta), (\gamma, \delta)) = [\mu(C) - \mu(A), \mu(B) + \mu(D)].$$

**PROOF.** We must show that  $\mu(2)$  can take on any value in  $[0, \mu(D)]$ . For any value of  $\mu(2)$  in  $[0, \mu(D)]$ , the Markov process with transition matrix

$$\begin{bmatrix} 1 - \alpha & \frac{\gamma\mu(1) - \beta\mu(4)}{\mu(1)} & \alpha - \gamma & \frac{\beta\mu(4)}{\mu(1)} \\ 0 & 1 - \delta & \delta & 0 \\ \beta - \gamma & \alpha \frac{\mu(2)}{\mu(3)} & 1 - \beta & \frac{\gamma\mu(3) - \alpha\mu(2)}{\mu(3)} \\ \delta & 0 & 0 & 1 - \delta \end{bmatrix}$$

gives  $\mu(2)$  that value (note that  $\mu(1) = \beta/(\alpha + \beta) - \mu(2)$ ,  $\mu(4) = \gamma/(\gamma + \delta) - \mu(2)$ ,  $\mu(3) = \mu(C) - \mu(A) + \mu(2)$ ). For this process  $A$  and  $B$  are balanced and respectful,  $C$  and  $D$  are level.

Note

$$\begin{aligned}\gamma\mu(1) - \beta\mu(4) &= \gamma\left(\frac{\beta}{\alpha + \beta} - \mu(2)\right) - \beta\left(\frac{\gamma}{\alpha + \beta} - \mu(2)\right) \\ &= (\beta - \gamma)\mu(2) \geq 0,\end{aligned}$$

so  $e_{12} \geq 0$ , and

$$\begin{aligned}\gamma\mu(3) - \alpha\mu(2) &= \gamma(\mu(C) - \mu(A) + \mu(2)) - \alpha\mu(2) \\ &= \gamma\left(\frac{\delta - \beta}{\alpha + \beta}\right) - (\alpha - \gamma)\mu(2) \\ &= (\alpha - \gamma)\left(\frac{\gamma}{\alpha + \beta} - \mu(2)\right) \\ &= (\alpha - \gamma)(\mu(D) - \mu(2)) \geq 0\end{aligned}$$

so  $e_{34} \geq 0$ . Thus all of the entries in the matrix are nonnegative. (If  $(\alpha, \beta) = (\gamma, \delta)$  and we set  $\mu(2) = 0$  or  $\mu(2) = \mu(D)$ , or if  $(\alpha, \beta) = (\delta, \gamma)$  and we set  $\mu(2) = \mu(D)$ , some of the entries in the matrix have the form  $0/0$ , but then the corresponding atoms have measure zero.) Q.E.D.

**LEMMA 2.7.** *If  $((\alpha, \beta), (\gamma, \delta))$  is in PFI but not in the special case, and  $\beta + \gamma > 1$ , then we cannot attain  $\mu(2) = 0$  by any Markov  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  (hence  $M((\alpha, \beta), (\gamma, \delta)) > \mu(C) - \mu(A)$ ).*

**PROOF.** Suppose we have  $\mu(2) = 0$  for a Markov  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$ . By Lemma 1.8 we can assume that  $C$  is level. Since  $\mu(2) = 0$ ,  $C$  is automatically uniform, so it is perfect. Now  $\mu(1) \neq 0$  and  $\mu(4) \neq 0$ , so by Lemma 2.4,  $A$  is not perfect. But since  $\mu(2) = 0$ ,  $A$  is level and if  $A$  were respectful, it would be uniform, hence perfect. Thus  $A$  is not respectful, whence by Lemmas 1.1 and 1.2,  $B$  must be level. Since  $B$  and  $C$  are level and  $\mu(2) = 0$ ,  $e_{31} + e_{34} = \beta + \gamma > 1$ , impossible since  $\mu(3) \neq 0$ . Q.E.D.

Lemma 2.7 together with Corollary 2.2 says that except for the pairs covered by Propositions 2.3 and 2.5 we cannot attain  $A \subseteq C$  (by a Markov process  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$ ). The following theorem is a strengthening of Lemma 2.7; it says that for the pairs in PFI, but not in the special case, the standard process cannot be improved upon (by a Markov process  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$ ).

**THEOREM 2.8.** *For  $((\alpha, \beta), (\gamma, \delta))$  in PFI but not in the special case,*

$$\mathfrak{M}((\alpha, \beta), (\gamma, \delta)) \cap [0, \text{St}((\alpha, \beta), (\gamma, \delta))] = \emptyset.$$

**PROOF.** By Proposition 2.3 we may assume  $\beta + \gamma > 1$ . We will assume we have an improvement on the standard process (a Markov  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathfrak{T}} < \text{St}((\alpha, \beta), (\gamma, \delta))$ ), and come to a contradiction.

Let

$$\begin{bmatrix} e_{11} & \cdots & e_{14} \\ \vdots & & \vdots \\ e_{41} & \cdots & e_{44} \end{bmatrix}$$

be the transition matrix for our improvement. By Lemma 1.8 we can assume that  $C$  is level. Let  $\gamma = \alpha - u$ ,  $\delta = \beta + w$  ( $w > u \geq 0$ ). Note

$$\mu(C) - \mu(A) = \frac{\beta + w}{\alpha + \beta + w - u} - \frac{\beta}{\alpha + \beta} = \frac{\alpha w + \beta u}{(\alpha + \beta)(\gamma + \delta)}.$$

If we have an improvement on the standard process, by equation (1.5),  $\mu(2)$  and  $\mu(3)$  must each have decreased by the same amount from the value they had for the standard process. Since for the standard process we have

$$\frac{\mu(2)}{\mu(3)} = \frac{[\mu(C) - \mu(A)](\beta + \gamma - 1)/(2 - (\beta + \gamma))}{[\mu(C) - \mu(A)](1/(2 - (\beta + \gamma)))} = \beta + \gamma - 1$$

for our improvement we have

$$\mu(2) < (\beta + \gamma - 1)\mu(3). \quad (2.2)$$

Hence

$$e_{32} \leq \frac{\mu(2)}{\mu(3)} < \beta + \gamma - 1. \quad (2.3)$$

Now if  $B$  were level, by (2.3) we would have

$$e_{31} + e_{32} + e_{34} = \beta + \gamma - e_{32} > \beta + \gamma - (\beta + \gamma - 1) = 1, \text{ impossible.}$$

Thus  $B$  is not level, so by Lemma 1.2  $B$  must be balanced, whence

$$e_{43} = \frac{\mu(3)}{\mu(4)}(e_{31} + e_{32} + e_{34} - \beta) = \frac{\mu(3)}{\mu(4)}(\gamma + e_{31} - \beta). \quad (2.4)$$

Also, since  $\mu(A \cap T(B)) = \beta\mu(B)$

$$e_{41} = \beta - e_{42} + \frac{\mu(3)}{\mu(4)}(\beta - e_{31} - e_{32}). \quad (2.5)$$

Suppose  $C$  is perfect. Then by Lemma 2.4,  $A$  cannot be perfect, and since  $B$  is not level, by the Basic Lemma we must have  $A$  and  $B$  respectful and balanced.

Then since  $B$  is respectful and  $C$  is perfect, the inflow into 2 equals

$$\mu(2)(\alpha + \delta - (\gamma + \delta)\mu(3)).$$

But by (2.2)

$$\begin{aligned}
\alpha + \delta - (\gamma + \delta) \mu(3) &= \alpha + \delta - (\gamma + \delta)(\mu(C) - \mu(A) + \mu(2)) \\
&= \gamma + \beta + w + u - \frac{\alpha w + \beta u}{\alpha + \beta} - (\gamma + \delta) \mu(2) \\
&> \gamma + \beta - (\gamma + \delta) \mu(2) \\
&= 1 + \frac{\mu(3)(\beta + \gamma - 1)}{\mu(3)} - (\gamma + \delta) \mu(2) \\
&> 1 + \left( \frac{1}{\mu(3)} - (\gamma + \delta) \right) \mu(2) \\
&> 1 + \left( \frac{1}{\mu(C)} - (\gamma + \delta) \right) \mu(2) \geq 1. \tag{2.6}
\end{aligned}$$

Impossible, since (by Lemma 2.7)  $\mu(2) > 0$ . Hence  $C$  is not perfect, so by Lemma 1.8, we can assume that  $D$  is level.

Thus, by (2.4) and (2.5)

$$\delta = e_{41} + e_{43} = \beta - e_{42} + \frac{\mu(3)}{\mu(4)} (\gamma - e_{32})$$

so

$$w = \frac{\mu(3)\gamma - \mu(3)e_{32} - \mu(4)e_{42}}{\mu(4)}.$$

Suppose  $B$  is respectful. Then

$$\begin{aligned}
w &= \frac{[\mu(C) - \mu(A) + \mu(2)]\gamma - \alpha\mu(2)}{\mu(D) - \mu(2)} \\
&= \frac{[(\alpha w + \beta u)/(\alpha + \beta)(\gamma + \delta) + \mu(2)]\gamma - \alpha\mu(2)}{\gamma/(\gamma + \delta) - \mu(2)}.
\end{aligned}$$

Hence

$$\frac{w\gamma}{\gamma + \delta} + \mu(2)(\alpha - \gamma - w) = \frac{(\alpha w + \beta u)\gamma}{(\alpha + \beta)(\gamma + \delta)}$$

or

$$\begin{aligned}
\mu(2)(u - w) &= \frac{\gamma}{\gamma + \delta} \left( \frac{(\alpha w + \beta u)}{\alpha + \beta} - w \right) \\
&= \left( \frac{\gamma}{\gamma + \delta} \right) \left( \frac{\beta}{\alpha + \beta} \right) (u - w).
\end{aligned}$$

Since  $u \neq w$  we get

$$\mu(2) = \left( \frac{\gamma}{\gamma + \delta} \right) \left( \frac{\beta}{\alpha + \beta} \right) = \mu(D) \mu(A).$$

Then  $\mu(2)$  has the same value as is attained by independently matching  $(\alpha, \beta)$  and  $(\gamma, \delta)$ , certainly no improvement over the standard process (since the independent matching is a simple process). Thus  $B$  is not respectful, and since  $B$  is not level, by the Basic Lemma  $A$  must be perfect.

Since  $A$  is perfect and  $D$  is level, by (2.6) we get

$$\begin{aligned} e_{21} + e_{23} + e_{24} &= \alpha + \delta - (\alpha + \beta) \mu(3) \\ &> \alpha + \delta - (\gamma + \delta) \mu(3) > 1. \end{aligned}$$

Impossible, since (by Lemma 2.7)  $\mu(2) > 0$ . Thus the theorem is proved.

**COROLLARY 2.9.** *For  $((\alpha, \beta), (\gamma, \delta))$  in PFI but not in the special case  $M((\alpha, \beta), (\gamma, \delta)) = \text{St}((\alpha, \beta), (\gamma, \delta))$ .*

**COROLLARY 2.10.** *If the proper form for  $((\alpha, \beta), (\delta, \gamma))$  is PFI, and  $((\alpha, \beta), (\gamma, \delta)) \notin \text{special case}$ , then*

$$\mathfrak{N}((\alpha, \beta), (\gamma, \delta)) \cap (1 - \text{St}((\alpha, \beta), (\delta, \gamma)), 1] \neq \emptyset.$$

**COROLLARY 2.11.**  *$M$  is not a metric; hence  $M \neq \bar{d}$ .*

**PROOF.**

$$M((.7, .8), (.6, .9)) = \mu(C) - \mu(A) = \frac{453}{6795},$$

$$M((.6, .9), (.6, .91)) = \text{St}((.6, .9), (.6, .91)) = \frac{54}{6795},$$

$$M((.7, .8), (.6, .91)) = \text{St}((.7, .8), (.6, .91)) = \frac{1099}{6795}.$$

Thus the Triangle Inequality fails.

**COROLLARY 2.12.**  *$M$  is not continuous as a function of  $\alpha, \beta, \gamma, \delta$ .*

**PROOF.**  $M$  is discontinuous at every  $((\alpha, \beta), (\alpha - u, \beta + u))$  with  $u > 0$  and  $\alpha + \beta - u > 1$ .

For example, for such pairs

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0, \epsilon \neq 0} M((\alpha, \beta), (\alpha - u, \beta + u + \epsilon)) \\ &= \frac{u}{\alpha + \beta} \left( \frac{\alpha + \beta - u}{2 + u - \alpha - \beta} \right) > \frac{u}{\alpha + \beta} = M((\alpha, \beta), (\alpha - u, \beta + u)). \end{aligned}$$

**3. Proper form II.** In §3 we determine  $\mathfrak{N}$  and  $M$  for all pairs in PFII. We find that  $M = \text{St}$  unless several inequalities hold, in which case  $M$  may be less than  $\text{St}$ . An example shows that the Triangle Inequality again fails, and also that  $\mathfrak{N}$  need not be an interval.

**THEOREM 3.1.** For  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$ , if  $\alpha + \delta \leq 1$ , or  $\beta + \gamma \leq 1$ ,

$$\mathfrak{M}((\alpha, \beta), (\gamma, \delta)) \cap [0, \text{St}((\alpha, \beta), (\gamma, \delta))] = \emptyset.$$

**PROOF.** We must show that there is no improvement on the standard process for  $((\alpha, \beta), (\gamma, \delta))$  (an improvement is a Markov  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathfrak{T}} < \text{St}((\alpha, \beta), (\gamma, \delta))$ ). There are three cases, F, G, and H.

**F:**  $\alpha + \delta \leq 1$  and  $\beta + \gamma \leq 1$ . The standard process for  $((\alpha, \beta), (\gamma, \delta))$  has transition matrix

$$\begin{bmatrix} 1 - \gamma & u & 0 & \alpha \\ \delta & 1 - (\alpha + \delta) & 0 & \alpha \\ \beta & 0 & 1 - (\beta + \gamma) & \gamma \\ \beta & 0 & w & 1 - \delta \end{bmatrix}$$

with  $\mu(2)(\alpha + \delta) = u\mu(1)$ , so for improvement need

$$\mu(2)(\alpha + \delta) < u\mu(1). \quad (3.1)$$

**G:**  $\alpha + \delta > 1$  and  $\beta + \gamma \leq 1$ . The standard process for  $((\alpha, \beta), (\gamma, \delta))$  has transition matrix

$$\begin{bmatrix} 1 - \gamma & u & 0 & \alpha \\ 1 - \alpha & 0 & \alpha + \delta - 1 & 1 - \delta \\ \beta & 0 & 1 - (\beta + \gamma) & \gamma \\ \beta & 0 & w & 1 - \delta \end{bmatrix}$$

with  $\mu(2) = u\mu(1)$ , so for improvement need

$$\mu(2) < u\mu(1). \quad (3.2)$$

**H:**  $\alpha + \delta \leq 1$  and  $\beta + \gamma > 1$ . The standard process for  $((\alpha, \beta), (\gamma, \delta))$  has transition matrix

$$\begin{bmatrix} 1 - \gamma & u & 0 & \alpha \\ \delta & 1 - (\alpha + \delta) & 0 & \alpha \\ 1 - \gamma & \beta + \gamma - 1 & 0 & 1 - \beta \\ \beta & 0 & w & 1 - \delta \end{bmatrix}$$

with  $\mu(3) = w\mu(4)$ , so for improvement need

$$\mu(3) < w\mu(4). \quad (3.3)$$

**Case F.** Assume we have an improvement on the standard process, so (3.1) holds. By Lemma 1.8 we can assume that  $A$  is level. Then by Lemma 1.4,  $e_{12} \geq \gamma - \alpha = u$ , so by (3.1)

$$\mu(1)e_{12} \geq u\mu(1) > (\alpha + \delta)\mu(2),$$

whence  $C$  is not respectful ( $C$  respectful requires  $\mu(1)e_{12} + \mu(3)e_{32} = \delta\mu(2)$ ). Since  $A$  is level, the outflow from 2 to  $B$  equals  $\alpha\mu(2)$ , and since the inflow into 2 is  $\geq \mu(1)e_{12} > (\alpha + \delta)\mu(2)$ ,  $e_{21} > \delta$ , so (since  $\mu(2) > 0$ )  $D$  is not level, whence  $C$  must be respectful. Contradiction.

*Case G.* Assume we have an improvement on the standard process, so (3.2) holds. By Lemma 1.8, we can assume that  $A$  is level. Then by Lemma 1.4,  $e_{12} \geq \gamma - \alpha = u$  so by (3.2)

$$\mu(1)e_{12} \geq u\mu(1) > \mu(2). \quad \text{Impossible.}$$

*Case H.* Assume we have an improvement on the standard process, so (3.3) holds. By Lemma 1.8 we can assume that  $B$  is level. Then by Lemma 1.4,  $e_{43} \geq \delta - \beta = w$  so by (3.3)

$$\mu(4)e_{43} \geq w\mu(4) > \mu(3). \quad \text{Impossible.} \quad \text{Q.E.D.}$$

**COROLLARY 3.2.** For  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$ , if  $\alpha + \delta \leq 1$  or  $\beta + \gamma \leq 1$ , then

$$M((\alpha, \beta), (\gamma, \delta)) = \text{St}((\alpha, \beta), (\gamma, \delta)).$$

**COROLLARY 3.3.** If the proper form for  $((\alpha, \beta), (\delta, \gamma))$  is  $((\alpha', \beta'), (\gamma', \delta')) \in \text{PFII}$  with  $\alpha' + \delta' \leq 1$  or  $\beta' + \gamma' \leq 1$ , then

$$\mathfrak{N}((\alpha, \beta), (\gamma, \delta)) \cap (1 - \text{St}((\alpha, \beta), (\delta, \gamma)), 1] = \emptyset.$$

**THEOREM 3.4.** If  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$  with  $\beta + \gamma > 1$  and  $\alpha + \delta > 1$ ,  $\gamma = \alpha + u$ ,  $\delta = \beta + w$ , then

$$\begin{aligned} & \mathfrak{N}((\alpha, \beta), (\gamma, \delta)) \cap [0, \text{St}((\alpha, \beta), (\gamma, \delta))) \\ &= \left[ \frac{2}{\alpha + \beta} \max\left(\frac{u\beta}{\delta + u}, \beta + \gamma - 1\right) + \frac{\alpha w - \beta u}{(\alpha + \beta)(\gamma + \delta)}, \right. \\ & \quad \left. \frac{2 - 2\delta}{\alpha + \beta} + \frac{\alpha w - \beta u}{(\alpha + \beta)(\gamma + \delta)} \right] \\ & \cap [0, \text{St}((\alpha, \beta), (\gamma, \delta))). \end{aligned}$$

**PROOF.** Note that

$$\mu(C) - \mu(A) = \frac{\beta + w}{\alpha + \beta + u + w} - \frac{\beta}{\alpha + \beta} = \frac{\alpha w - \beta u}{(\alpha + \beta)(\gamma + \delta)}. \quad (3.4)$$

The standard process has transition matrix

$$\begin{bmatrix} 1 - \gamma & u & 0 & \alpha \\ 1 - \alpha & 0 & \alpha + \delta - 1 & 1 - \delta \\ 1 - \gamma & \beta + \gamma - 1 & 0 & 1 - \beta \\ \beta & 0 & w & 1 - \delta \end{bmatrix}.$$

For the standard process  $\mu(2) = u\mu(1) + (\beta + \gamma - 1)\mu(3)$ . If we have an improvement on the standard process (a Markov  $\mathcal{T} \in Y((\alpha, \beta), (\gamma, \delta))$  with  $d_{\mathcal{T}} < \text{St}((\alpha, \beta), (\gamma, \delta))$ ), we have

$$\mu(2) < u\mu(1) + (\beta + \gamma - 1)\mu(3). \quad (3.5)$$

We will assume that (3.5) holds and see what this implies. By Lemma 1.8 we can assume that  $C$  is level for the improvement.

First we will show that  $A$  cannot also be level.

Suppose  $A$  is level. Then  $e_{12} \geq u$ , so by (3.5),  $e_{32} < \beta + \gamma - 1$ . Since  $C$  is level,

$$e_{34} = \gamma - e_{32} > \gamma - (\beta + \gamma - 1) = 1 - \beta,$$

thus  $e_{31} + e_{32} < \beta$  so  $B$  is not level. Since  $A$  and  $C$  are assumed to be level

$$e_{12} = u + e_{13}. \quad (3.6)$$

If  $A$  is balanced, then (since  $A$  is level)

$$e_{21}\mu(2) = e_{12}\mu(1); \quad (3.7)$$

hence by (3.5), (3.6) and (3.7)

$$e_{13}\mu(1) + e_{23}\mu(2) < \beta\mu(3)$$

whence  $A$  is not respectful. Thus  $A$  is not both balanced and respectful, and since  $B$  is not level, by the Basic Lemma we must have  $A$  perfect. Thus

$$e_{12} + e_{13} + e_{14} = \gamma + e_{13} = \gamma + (\alpha + \beta)\mu(3)$$

whence  $(\alpha + \beta)\mu(3) \leq 1 - \gamma$ , and since  $\mu(2) \leq \mu(3)$ ,

$$(\alpha + \beta)\mu(2) \leq 1 - \gamma. \quad (3.8)$$

But by (3.6)

$$\mu(2) \geq e_{12}\mu(1) = (u + e_{13})\mu(1) = u\mu(1) + (\alpha + \beta)\mu(1)\mu(3)$$

whence, by (3.5),  $(\alpha + \beta)\mu(1) < \beta + \gamma - 1$ , so

$$(\alpha + \beta)\mu(2) = \beta - (\alpha + \beta)\mu(1) > \beta - (\beta + \gamma - 1) = 1 - \gamma,$$

direct contradiction to (3.8). Thus we cannot have an improvement on the standard process with  $A$  (and  $C$ ) level.

Thus if we have an improvement on the standard process (with  $C$  level),  $A$  cannot be level, so  $A$  must be balanced. Then

$$\begin{aligned} \mu(2)e_{21} &= \mu(1)(e_{12} + e_{13} + e_{14} - \alpha) \\ &= \mu(1)(\gamma + e_{13} - \alpha) = \mu(1)(u + e_{13}). \end{aligned} \quad (3.9)$$

Thus

$$\mu(2)(e_{21} + e_{23}) = u\mu(1) + \mu(1)e_{13} + \mu(2)e_{23},$$

so by (3.5),

$$\mu(1)e_{13} + \mu(2)e_{23} < (\beta + \gamma - 1)\mu(3) \leq \beta\mu(3).$$

Thus  $A$  is not respectful. Since  $A$  is not level either, by the Basic Lemma  $B$  must be perfect. Then, since  $\mu(2) \neq 0$  and  $\mu(3) \neq 0$ , by Lemma 2.4,  $C$  cannot be perfect, whence by Lemma 1.8 we can assume that  $D$  is level.

*Summarizing.* Any distance less than  $\text{St}((\alpha, \beta), (\gamma, \delta))$  which can be attained by a Markov process in  $Y((\alpha, \beta), (\gamma, \delta))$  for  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$  with  $\alpha + \delta > 1$  and  $\beta + \gamma > 1$  can be attained by a Markov process in  $Y((\alpha, \beta), (\gamma, \delta))$  with  $A$  balanced,  $B$  perfect,  $C$  and  $D$  level.

Assume we have such a process. Modulo three degrees of freedom, our process is determined. We will let our parameters be  $e_{12}$ ,  $e_{13}$  and  $\mu(2)$ .

Since  $B$  is perfect,

$$\begin{aligned} e_{31} = e_{41} &= (\alpha + \beta) \mu(1) = \beta - (\alpha + \beta) \mu(2), \\ e_{32} = e_{42} &= (\alpha + \beta) \mu(2). \end{aligned}$$

Since  $A$  is balanced and  $C$  is level,

$$e_{21} = \frac{\mu(1)}{\mu(2)} (u + e_{13}).$$

Since  $C$  is level and  $B$  is perfect,

$$e_{34} = \gamma - (\alpha + \beta) \mu(2), \quad e_{14} = \gamma - e_{12}.$$

Since  $D$  is level and  $B$  is perfect,

$$\begin{aligned} e_{23} &= \delta - e_{21} = \delta - \frac{\mu(1)}{\mu(2)} (u + e_{13}), \\ e_{43} &= \delta - (\alpha + \beta) \mu(1) = w + (\alpha + \beta) \mu(2). \end{aligned}$$

Since measure entering 2 ( $\alpha\mu(2) + e_{12}\mu(1)$ ) equals measure leaving 2 ( $\delta\mu(2) + e_{24}\mu(2)$ ),

$$e_{24} = \alpha - \delta + \frac{\mu(1)}{\mu(2)} e_{12}.$$

Note

$$\begin{aligned} e_{11} &= 1 - e_{12} - e_{13} - e_{14} = 1 - \gamma - e_{13}, \\ e_{22} &= 1 - e_{21} - e_{23} - e_{24} = 1 - \alpha - \frac{\mu(1)}{\mu(2)} e_{12}, \\ e_{33} &= 1 - e_{31} - e_{32} - e_{34} = 1 - \beta - \gamma + (\alpha + \beta) \mu(2), \\ e_{44} &= 1 - e_{41} - e_{42} - e_{43} = 1 - \delta - (\alpha + \beta) \mu(2). \end{aligned}$$

Hence the transition matrix for the improvement is

$$\begin{pmatrix} 1 - \gamma - e_{13} & e_{12} & e_{13} & \gamma - e_{12} \\ \frac{\mu(1)}{\mu(2)}(u + e_{13}) & 1 - \alpha - \frac{\mu(1)}{\mu(2)}e_{12} & \delta - \frac{\mu(1)}{\mu(2)}(u + e_{13}) & \alpha - \delta + \frac{\mu(1)}{\mu(2)}e_{12} \\ \beta - (\alpha + \beta)\mu(2) & (\alpha + \beta)\mu(2) & (\alpha + \beta)\mu(2) - (\beta + \gamma - 1) & \gamma - (\alpha + \beta)\mu(2) \\ \beta - (\alpha + \beta)\mu(2) & (\alpha + \beta)\mu(2) & w + (\alpha + \beta)\mu(2) & 1 - \delta - (\alpha + \beta)\mu(2) \end{pmatrix}.$$

We will now investigate for which values of  $\mu(2)$  we can choose  $e_{12}$  and  $e_{13}$  so that all entries in the matrix are nonnegative; these values, the ones for which the following inequalities hold, are the values of  $\mu(2)$  attainable with  $A$  balanced,  $B$  perfect,  $C$  and  $D$  level.

$$e_{13} \leq 1 - \gamma \quad (e_{11} \geq 0), \quad (3.10)$$

$$e_{12} \leq \gamma \quad (e_{14} \geq 0), \quad (3.11)$$

$$e_{12}\mu(1) \leq (1 - \alpha)\mu(2) \quad (e_{22} \geq 0), \quad (3.12)$$

$$(u + e_{13})\mu(1) \leq \delta\mu(2) \quad (e_{23} \geq 0), \quad (3.13)$$

$$(\delta - \alpha)\mu(2) \leq e_{12}\mu(1) \quad (e_{24} \geq 0), \quad (3.14)$$

$$(\alpha + \beta)\mu(2) \leq \beta \quad (e_{31} \geq 0), \quad (3.15)$$

$$\beta + \gamma - 1 \leq (\alpha + \beta)\mu(2) \quad (e_{33} \geq 0), \quad (3.16)$$

$$(\alpha + \beta)\mu(2) \leq \gamma \quad (e_{34} \geq 0), \quad (3.17)$$

$$(\alpha + \beta)\mu(2) \leq 1 - \delta \quad (e_{44} \geq 0). \quad (3.18)$$

*Observations.* A necessary condition to satisfy (3.11) and (3.14) is that

$$\gamma\mu(1) \geq (\delta - \alpha)\mu(2),$$

$$\gamma \left( \frac{\beta}{\alpha + \beta} - \mu(2) \right) \geq (\delta - \alpha)\mu(2),$$

$$\frac{\gamma\beta}{\alpha + \beta} \geq (\delta - \alpha + \gamma)\mu(2) = (\delta + u)\mu(2),$$

or

$$\mu(2) \leq \frac{\gamma\beta}{(\delta + u)(\alpha + \beta)}. \quad (3.19)$$

Furthermore, this is sufficient to satisfy (3.11), (3.14), and (3.12) since if (3.19) holds, we can set

$$e_{12} = \begin{cases} 0 & \text{if } (\delta - \alpha) \leq 0, \\ \frac{(\delta - \alpha)\mu(2)}{\mu(1)} & \text{if } (\delta - \alpha) > 0. \end{cases}$$

Notice that (3.18)  $\Rightarrow$  (3.17), since  $1 - \delta < \gamma$ .

To satisfy (3.13) it is necessary that  $\delta\mu(2) \geq u\mu(1)$  (we need  $\mu(1)e_{13} \geq 0$ ),

and

$$\delta\mu(2) \geq u\mu(1) = u\left(\frac{\beta}{\alpha + \beta} - \mu(2)\right)$$

iff

$$\mu(2) \geq \frac{u\beta}{(\delta + u)(\alpha + \beta)}. \quad (3.20)$$

Furthermore, this is sufficient to satisfy (3.10) and (3.13) since we can set  $e_{13} = 0$ .

Thus, by (3.15), (3.16), (3.18), (3.19), (3.20) and the preceding observations, the values of  $\mu(2)$  attainable with  $A$  balanced,  $B$  perfect,  $C$  and  $D$  level are precisely those in the interval

$$\left[ \max\left(\frac{u\beta}{(\delta + u)(\alpha + \beta)}, \frac{\beta + \gamma - 1}{\alpha + \beta}\right), \min\left(\frac{\gamma\beta}{(\delta + u)(\alpha + \beta)}, \frac{1 - \delta}{\alpha + \beta}, \frac{\beta}{\alpha + \beta}\right) \right]. \quad (3.21)$$

Since independently matching  $(\alpha, \beta)$  and  $(\gamma, \delta)$  is a simple process, the value given  $\mu(2)$  by the standard process for  $((\alpha, \beta), (\gamma, \delta))$  is  $\leq \mu(A)\mu(D)$ . But

$$\mu(A)\mu(D) < \min\left(\frac{\gamma\beta}{(\delta + u)(\alpha + \beta)}, \frac{\beta}{\alpha + \beta}\right). \quad (3.22)$$

By (1.5), (3.4), (3.21), (3.22) and the preceding arguments, the theorem follows.

**COROLLARY 3.5.** *If  $((\alpha, \beta), (\gamma, \delta)) \in \text{PFII}$ ,  $\gamma = \alpha + u$ ,  $\delta = \beta + w$ ,  $\beta + \gamma > 1$ ,  $\alpha + \delta > 1$ , then*

$$M((\alpha, \beta), (\gamma, \delta))$$

$$= \begin{cases} \min\left(\text{St}((\alpha, \beta), (\gamma, \delta)), \frac{2}{\alpha + \beta} \max\left(\frac{u\beta}{\delta + u}, \beta + \gamma - 1\right) + \frac{\alpha w - \beta u}{(\alpha + \beta)(\gamma + \delta)}\right) & \text{if } \max\left(\frac{u\beta}{\delta + u}, \beta + \gamma - 1\right) \leq 1 - \delta, \\ \text{St}((\alpha, \beta), (\gamma, \delta)) & \text{otherwise.} \end{cases}$$

**COROLLARY 3.6.** *If the proper form for  $((\alpha, \beta), (\delta, \gamma))$  is  $((\alpha', \beta'), (\gamma', \delta')) \in \text{PFII}$ ,  $\gamma' = \alpha' + u'$ ,  $\delta' = \beta' + w'$ ,  $\beta' + \gamma' > 1$ ,  $\alpha' + \delta' > 1$ , then*

$$\begin{aligned}
& \mathfrak{N}((\alpha, \beta), (\gamma, \delta)) \cap (1 - \text{St}((\alpha, \beta), (\delta, \gamma)), 1] \\
&= \left[ 1 - \frac{2 - 2\delta'}{\alpha' + \beta'} - \frac{\alpha'w' - \beta'u'}{(\alpha + \beta)(\gamma + \delta)}, \right. \\
&\quad \left. 1 - \frac{2}{\alpha' + \beta'} \max\left(\frac{u'\beta'}{\delta' + u'}, \beta' + \gamma' - 1\right) - \frac{\alpha'w' - \beta'u'}{(\alpha + \beta)(\gamma + \delta)} \right] \\
&\cap (1 - \text{St}((\alpha, \beta), (\delta, \gamma)), 1].
\end{aligned}$$

Note that in some cases of Theorem 3.4 we will have

$$\mathfrak{N}((\alpha, \beta), (\gamma, \delta)) \cap [0, \text{St}((\alpha, \beta), (\gamma, \delta))) \neq \emptyset.$$

EXAMPLE 1. Let  $(\alpha, \beta) = (.5, .44)$ ,  $(\gamma, \delta) = (.58, .966)$ . Then

$$\begin{aligned}
\max\left(\frac{u\beta}{(\delta + u)(\alpha + \beta)}, \frac{\beta + \gamma - 1}{\alpha + \beta}\right) &= \frac{u\beta}{(\delta + u)(\alpha + \beta)} \\
&= \frac{(.08)(.44)}{(1.046)(.94)} \approx .0358000, \\
\min\left(\frac{\gamma\beta}{(\delta + u)(\alpha + \beta)}, \frac{1 - \delta}{\alpha + \beta}, \frac{\beta}{\alpha + \beta}\right) &= \frac{1 - \delta}{\alpha + \beta} \\
&= \frac{.034}{.94} \approx .0361702.
\end{aligned}$$

Hence by (3.21) we can attain

$$\mu(2) = \frac{(.08)(.44)}{(1.046)(.94)} \approx .0358000.$$

The standard process for this pair has

$$\begin{aligned}
\mu(2) &= \frac{u\delta + w\gamma(\alpha + \beta - 1)}{(\alpha + \beta)(\gamma + \delta)(2 - (\alpha + \beta))} \\
&= \frac{(.08)(.966) + (.526)(.58)(-.06)}{(.94)(1.546)(1.06)} \approx .0382848.
\end{aligned}$$

Since we can lower  $\mu(2)$  to  $\approx .0358000$ , we can improve upon the standard process

Further note that  $\mathfrak{N}((.5, .44), (.58, .966))$  is not an interval.

The values of  $\mu(2)$  less than that for the standard process which can be attained by a Markov  $\mathfrak{T} \in Y((\alpha, \beta), (\gamma, \delta))$  are precisely those in

$$[u\beta / (\delta + u)(\alpha + \beta), (1 - \delta) / (\alpha + \beta)],$$

and this interval is entirely disjoint from the range of values of  $\mu(2)$  which are attained by the simple processes for  $((.5, .44), (.58, .966))$  (Lemma 1.5). Hence, since  $\mu(2 \cup 3) = 2\mu(2) + \mu(C) - \mu(A)$ ,  $\mathfrak{M}((.5, .44), (.58), (.966))$  contains two disjoint nonempty closed intervals and none of the values between those intervals. Thus we have shown:

**COROLLARY 3.7.**  $\mathfrak{M}((\alpha, \beta), (\gamma, \delta))$  need not form an interval.

**EXAMPLE 2.** If in the preceding example we hold  $\alpha, \beta$ , and  $\delta$  fixed and increase  $\gamma$  until

$$\frac{u\beta}{(\delta + u)(\alpha + \beta)} = \frac{1 - \delta}{\alpha + \beta}$$

$$\left( \gamma = \alpha + \frac{(\delta)(1 - \delta)}{(\beta + \delta - 1)} = .5 + \frac{(.966)(.034)}{.406} \approx .5808966 \right),$$

then  $\mathfrak{M}((\alpha, \beta), (\gamma, \delta))$  will consist of those values attained by the simple processes for  $((\alpha, \beta), (\gamma, \delta))$  plus the single additional value

$$\frac{2(1 - \delta)}{\alpha + \beta} + \frac{\alpha w - \beta u}{(\alpha + \beta)(\gamma + \delta)} \quad (= M((\alpha, \beta), (\gamma, \delta))).$$

By examining

$$\lim_{\epsilon \downarrow 0} M\left((.5, .44), \left(.5 + \frac{(.966)(.034)}{.406} + \epsilon, .966\right)\right),$$

we could again get the results of Corollaries 2.11 and 2.12.

Lemma 1.5, Proposition 2.6, Theorems 2.8, 3.1, and 3.4, and Corollaries 2.10, 3.3, and 3.6 compute  $\mathfrak{M}$ ; Propositions 2.3 and 2.5, and Corollaries 2.9, 3.2, and 3.5 compute  $M$ . We would like to know  $\bar{d}$ .

**Appendix.** (1) In the text we found that  $M$  does not equal  $\bar{d}$  since  $M$  fails to satisfy the Triangle Inequality. Consider  $M$ 's associated metric,  $M^*$ :

$$M^*((\alpha, \beta), (\gamma, \delta)) = \inf_{n \in \mathbb{N}} \left\{ \sum_{i=0}^n M((\alpha_i, \beta_i), (\alpha_{i+1}, \beta_{i+1})): (\alpha_0, \beta_0) = (\alpha, \beta) \right. \\ \left. \text{and } (\alpha_{n+1}, \beta_{n+1}) = (\gamma, \delta) \right\}.$$

**THEOREM.** If  $\bar{d}((\alpha, \beta), (\gamma, \delta)) < M((\alpha, \beta), (\gamma, \delta))$ , then  $\bar{d}((\alpha, \beta), (\gamma, \delta)) < M^*((\alpha, \beta), (\gamma, \delta))$ .

Hence  $M^*$  never attains  $\bar{d}$  unless  $M$  already attains  $\bar{d}$ .

(2) In the text the only method of showing that  $\bar{d}((\alpha, \beta), (\gamma, \delta)) \neq M((\alpha, \beta), (\gamma, \delta))$  was finding a process  $(\alpha', \beta')$  such that

$$M((\alpha, \beta), (\alpha', \beta')) + M((\alpha', \beta'), (\gamma, \delta)) < M((\alpha, \beta), (\gamma, \delta)).$$

In [2] it is shown that for  $\alpha < \frac{1}{2}$ ,  $\bar{d}((\alpha, \alpha), (1 - \alpha, 1 - \alpha)) = (1 - 2\alpha)/2$ . Since it can quickly be shown that for such pairs  $M$  is  $(1 - 2\alpha)/2(1 - \alpha)$ , this gives a short proof that  $M \neq \bar{d}$ ; note that this proof does not involve showing that  $M$  fails the Triangle Inequality.

(3) We have recently found [3] that  $\bar{d}((\alpha, \beta), (\gamma, \delta)) < M((\alpha, \beta), (\gamma, \delta))$  except for the cases covered by Propositions 2.3 and 2.5. Thus, unless  $M((\alpha, \beta), (\gamma, \delta))$  is the partition distance, it is greater than  $\bar{d}((\alpha, \beta), (\gamma, \delta))$ . Between 3-state Markov processes with positive entropy  $M$  can sometimes equal  $\bar{d}$  and not be the partition distance.

(4) Kamae observed that if  $\bar{d}((\alpha, \beta), (\gamma, \delta))$  is the partition distance, then

$$\forall n \in \mathbb{N}^+, \quad ((\alpha, \beta)^n, (\gamma, \delta)^n) \text{ has an equivalent in PFI} \quad (*)$$

and he conjectured  $(*)$  suffices to insure that  $\bar{d}((\alpha, \beta), (\gamma, \delta))$  is the partition distance. In [4] it is shown that for  $((\alpha, \beta), (\gamma, \delta))$  in PFI,  $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$  whenever (i)  $\alpha + \beta + \gamma \leq 2$  and  $w(\beta + \gamma - 1) \leq u\beta$ , or (ii)  $\beta + \gamma + \delta \leq 2$  and  $u(\beta + \gamma - 1) \leq w\beta$ , or (iii)  $\alpha + \beta = \gamma + \delta$ . The construction does not settle Kamae's conjecture, however, for not all pairs satisfying  $(*)$  satisfy (i) or (ii) or (iii).

## REFERENCES

1. K. R. Berg, *Convolution of invariant measures, maximal entropy*, Math. Systems Theory 3 (1969), 146–150. MR 40 #1582.
2. M. H. Ellis, *The  $\bar{d}$ -distance between two Markov processes cannot always be attained by a Markov joining*, Israel J. Math. 24 (1976), 269–273. MR 54 #2912.
3. ———, *Conditions for attaining  $\bar{d}$  by a Markovian joining* (submitted).
4. ———, *On a conjecture by Kamae* (submitted).
5. H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory 1 (1967), 1–49. MR 35 #4369.
6. T. Kamae, U. Krengel and G. L. O'Brien, *Stochastic inequalities on partially ordered spaces*, Ann. of Probability 5 (1977), 899–912.
7. J. G. Kemeny and J. L. Snell, *Finite Markov chains*, Van Nostrand, Princeton, N. J., 1960. MR 22 #5998.
8. D. S. Ornstein, *An application of ergodic theory to probability theory*, Ann. of Probability 1 (1973), 43–65. MR 50 #1326
9. ———, *Ergodic theory, randomness and dynamical systems*, Yale Univ. Press, New Haven, Conn., 1974.

DEPARTMENT OF MATHEMATICS, STATE UNIVERSITY OF NEW YORK AT ALBANY, ALBANY, NEW YORK 12222

*Current address:* Department of Mathematics, University of British Columbia, Vancouver, B. C., Canada