The Annals of Probability
1986, Vol. 14, No. 3, 873–876

ON MAXIMAL AND DISTRIBUTIONAL COUPLING[1]

BY HERMANN THORISSON

University of Göteborg

A simple construction of maximal coupling is given by way of a distributional coupling concept.

1. Introduction. Let $Z = (Z_t)_0^\infty$ and $Z' = (Z'_t)_0^\infty$ be discrete time stochastic processes on a general state space $(E, \mathscr{E})$. It is known that under topological restrictions on $(E, \mathscr{E})$ there exists a maximal coupling of Z and Z'; see Griffeath [2], [3], Pitman [4], and Goldstein [1].

In the author's paper [5] a distributional coupling concept was introduced to deal with the difficulties arising when certain continuous time stochastic processes on a general state space are coupled. It is the purpose of the present note to show how this distributional approach allows a short and transparent construction of a maximal coupling, without any restrictions on $(E, \mathscr{E})$. The nondistributional maximal coupling result is then a direct corollary to the distributional one. We restrict our attention to discrete time processes, mentioning the continuous time problem only in a remark.

For further information on coupling we refer the reader to the papers quoted above.

2. Nondistributional coupling. For nonnegative integer valued random variables T put $\theta_T Z = (Z_{T+t})_0^\infty$ on $\{T < \infty\}$ and $\theta_T Z = (z, z, \dots)$ on $\{T = \infty\}$, where z is some fixed element of E. Let $=_D$ denote identity in distribution and $\|\cdot\|$ the total variation norm.

The following is a version of the traditional definition of coupling:

DEFINITION 1. $\hat{Z} = (\hat{Z}_t)_0^\infty$ and $\hat{Z}' = (\hat{Z}'_t)_0^\infty$ is a *nondistributional coupling* of Z and Z' with *coupling epoch* T if

(a) $\hat{Z} =_D Z$ and $\hat{Z}' =_D Z'$,
(b) $\theta_T \hat{Z} = \theta_T \hat{Z}'$.

The coupling is called *successful* if $\mathbb{P}(T < \infty) = 1$. In this case the following inequality yields a strong limit result as $n \to \infty$, which explains much of the interest in coupling:

THE COUPLING INEQUALITY. *For any integer* $n \geq 0$

$$\|\mathbb{P}(\theta_n Z \in \cdot) - \mathbb{P}(\theta_n Z' \in \cdot)\| \leq 2\mathbb{P}(T > n).$$

Received November 1984.

[1]Supported by the Swedish Natural Science Research Council and by the Icelandic Science Foundation.

AMS 1980 *subject classification*. Primary 60G05.

Key words and phrases. Coupling, maximal coupling, distributional coupling.

PROOF. Due to (b), $\mathbb{P}(\theta_n\hat{Z} \in A, T \le n) = \mathbb{P}(\theta_n\hat{Z}' \in A, T \le n)$ and thus

$$\begin{aligned}\mathbb{P}(\theta_n\hat{Z} \in A) - \mathbb{P}(\theta_n\hat{Z}' \in A) &= \mathbb{P}(\theta_n\hat{Z} \in A, T > n) - \mathbb{P}(\theta_n\hat{Z}' \in A, T > n)\\ &\le \mathbb{P}(\theta_n\hat{Z} \in A, T > n) \le \mathbb{P}(T > n).\end{aligned}$$

Take supremum in $A \in \mathscr{E}^{\{0,1,\dots\}}$, multiply by 2 and apply (a) to obtain the desired result. □

3. Distributional coupling.

We now weaken the requirement that $\hat{Z}$ and $\hat{Z}'$ ultimately coincide and demand rather that $\hat{Z}$ behaves probabilistically from a time T onwards as $\hat{Z}'$ does from a time T' onwards, or more precisely:

DEFINITION 2. $\hat{Z}$ and $\hat{Z}'$ is a *distributional coupling* of Z and Z' with *coupling epochs* T and T' if (a) of Definition 1 holds and

(b′) $(\theta_T\hat{Z}, T) =_D (\theta_{T'}\hat{Z}', T')$.

With this coupling concept *the coupling inequality* still holds: In the proof apply (b′) instead of (b) and replace T by T' in the appropriate places.

REMARK 1. If there exists a regular version of the conditional distribution of $\hat{Z}$ given $(\theta_T\hat{Z}, T)$—in particular, if E is Polish and $\mathscr{E}$ its Borel subsets—then $\hat{Z}$ and $\hat{Z}'$ can be glued together through the identically distributed random elements in (b′) transforming $=_D$ into $=$ (see Construction 1.1 in [5]). Thus in this case the distributional coupling in Definition 2 can be made nondistributional.

REMARK 2. Since T is discrete there certainly exists a regular version of the conditional distribution of $\hat{Z}$ given T. Thus we may assume that $T = T'$.

4. Maximal coupling.

DEFINITION 3. The coupling (distributional or nondistributional) is *maximal* if *the coupling inequality* is an identity:

$$\|\mathbb{P}(\theta_n Z \in \cdot\,) - \mathbb{P}(\theta_n Z' \in \cdot\,)\| = 2\mathbb{P}(T > n), \qquad n \ge 0.$$

THEOREM. *There exists a maximal distributional coupling.*

The theorem is proved in the final two sections.

As an immediate consequence (see Remark 1 above) we have the following:

COROLLARY. *If E is Polish and $\mathscr{E}$ its Borel subsets then there exists a maximal nondistributional coupling.*

REMARK 3. For continuous time processes $Z = (Z_s)_{[0,\infty)}$, $Z' = (Z'_s)_{[0,\infty)}$, there does not in general exist a maximal distributional coupling such that

$$\|\mathbb{P}(\theta_t Z \in \cdot\,) - \mathbb{P}(\theta_t Z' \in \cdot\,)\| = 2\mathbb{P}(T > t), \qquad t \in [0, \infty),$$

since although the left-hand side is nonincreasing in t it need not be right- or left-continuous. However, if we apply the theorem to $Y = (Y_t)_0^\infty$ and $Y' = (Y'_t)_0^\infty$ where $Y_t = (Z_{td+s})_{[0,d)}$, $Y'_t = (Z'_{td+s})_{[0,d)}$ with d a positive constant, then we obtain the following result: *For any* $d \in (0, \infty)$ *there exists a distributional coupling of* Z *and* Z' *maximal at all* $t \in \{0, d, 2d, \dots\}$. In particular, this means that also in the continuous time case *total variation convergence is equivalent to the existence of a successful distributional coupling*. Observe, that even if $(E, \mathscr{E})$ is Polish the state space of Y and Y', $(E^{[0,d)}, \mathscr{E}^{[0,d)})$, is not. Thus we need path regularity conditions (such as right- or left-continuity) in order to get the regular conditional distributions needed to establish the analogous nondistributional result.

5. Proof of the theorem. The following lemma is the key part of the proof:

LEMMA. *Let* $\mathscr{F}_0 \supset \mathscr{F}_1 \supset \cdots$ *be a nonincreasing sequence of* σ-*algebras and* π, π' *two probability measures on* $\mathscr{F}_0$. *Let* $(\pi)_{\mathscr{F}_n}$ *denote the restriction of* π *to* $\mathscr{F}_n$ *and* $\pi \wedge \pi'$ *the greatest common component of* π *and* π'. *Then there exist subprobability measures* λ_n, λ'_n, $0 \le n \le \infty$, *on* $\mathscr{F}_0$ *such that*

$$\lambda_0 + \cdots + \lambda_\infty = \pi \quad \text{and} \quad \lambda'_0 + \cdots + \lambda'_\infty = \pi', \tag{1}$$

$$(\lambda_0 + \cdots + \lambda_n)_{\mathscr{F}_n} = (\lambda'_0 + \cdots + \lambda'_n)_{\mathscr{F}_n} = (\pi)_{\mathscr{F}_n} \wedge (\pi')_{\mathscr{F}_n}, \qquad 0 \le n < \infty. \tag{2}$$

In order to prove the theorem it is no restriction to let Z and Z' both be defined on sequence space, $E^{\{0,1,\dots\}}$. Put $\mathscr{F}_n = \sigma\{\theta_n Z\} = \sigma\{\theta_n Z'\}$ and let π, π' be the distributions of Z, Z', respectively. Define probability measures λ, λ' on $(E^{\{0,1,\dots\}} \times \{0, \dots, \infty\}, \mathscr{F}_0 \otimes B\{0, \dots, \infty\})$ by $\lambda(\cdot \times \{n\}) = \lambda_n$, $\lambda'(\cdot \times \{n\}) = \lambda'_n$ and let $(\hat{Z}, T), (\hat{Z}', T')$ be governed by λ, λ', respectively. Clearly (1) and the first equality in (2) imply the conditions of Definition 2. Further, (2) yields the final equality in

$$\begin{aligned} \|\mathbb{P}(\theta_n \hat{Z} \in \cdot) - \mathbb{P}(\theta_n \hat{Z}' \in \cdot)\| &= \|(\pi)_{\mathscr{F}_n} - (\pi')_{\mathscr{F}_n}\| \\ &= 2\big(1 - \|(\pi)_{\mathscr{F}_n} \wedge (\pi')_{\mathscr{F}_n}\|\big) \\ &= 2\big(1 - \|\lambda_0 + \cdots + \lambda_n\|\big), \end{aligned}$$

proving maximality since $\|\lambda_0 + \cdots + \lambda_n\| = \lambda(E^{\{0,1,\dots\}} \times \{0, \dots, n\}) = \mathbb{P}(T \le n)$.

6. Proof of the lemma. We start by defining the subprobability measures λ_n, λ'_n on $\mathscr{F}_n$ such that (2) holds:

$$\lambda_0 = \lambda'_0 = \pi \wedge \pi',$$

$$(\lambda_n)_{\mathscr{F}_n} = (\lambda'_n)_{\mathscr{F}_n} = (\pi)_{\mathscr{F}_n} \wedge (\pi')_{\mathscr{F}_n} - \big((\pi)_{\mathscr{F}_{n-1}} \wedge (\pi')_{\mathscr{F}_{n-1}}\big)_{\mathscr{F}_n}, \qquad 1 \le n < \infty.$$

Make the induction assumption that λ_k can be extended as a subprobability measure from $\mathscr{F}_k$ to $\mathscr{F}_0$ for $0 \le k < n$ and that $\lambda_0 + \cdots + \lambda_{n-1} \le \pi$ (this certainly holds for $n = 1$).

Extend λ_n from $\mathscr{F}_n$ to $\mathscr{F}_0$ by

$$(3) \qquad \lambda_n(A) = \int (\pi - \lambda_0 - \cdots - \lambda_{n-1})(A|\mathscr{F}_n)\, d(\lambda_n)_{\mathscr{F}_n}, \qquad A \in \mathscr{F}_0.$$

By (2)

$$(4) \qquad (\lambda_n)_{\mathscr{F}_n} \le (\pi - \lambda_0 - \cdots - \lambda_{n-1})_{\mathscr{F}_n},$$

implying that the set-function λ_n is independent of the version of $(\pi - \lambda_0 - \cdots - \lambda_{n-1})(\cdot|\mathscr{F}_n)$. Thus for a *given* sequence of disjoint events in $\mathscr{F}_0$ we can choose a version that is σ-additive for that *specific* sequence. Hence λ_n is σ-additive. Further, λ_n is nonnegative and from (3) and (4) we obtain $\lambda_0 + \cdots + \lambda_n \le \pi$.

By induction we have proved that λ_n, $0 \le n < \infty$, can be extended as subprobability measures to $\mathscr{F}_0$ such that $\lambda_0 + \lambda_1 + \cdots \le \pi$. By symmetry the same holds with λ_n replaced by λ'_n and π by π'. Define

$$\lambda_\infty = \pi - \lambda_0 - \lambda_1 - \cdots \quad \text{and} \quad \lambda'_\infty = \pi' - \lambda'_0 - \lambda'_1 - \cdots$$

to obtain (1) and complete the proof.

REFERENCES

[1] GOLDSTEIN, S. (1979). Maximal coupling. *Z. Wahrsch. verw. Gebiete* **46** 193–204.

[2] GRIFFEATH, D. (1975). A maximal coupling for Markov chains. *Z. Wahrsch. verw. Gebiete* **31** 95–106.

[3] GRIFFEATH, D. (1978). Coupling methods for Markov processes. In *Advances in Mathematics Supplementary Studies: Studies in Probability and Ergodic Theory* **2** 1–43. Academic, New York.

[4] PITMAN, J. W. (1976). On coupling of Markov chains. *Z. Wahrsch. verw. Gebiete* **35** 315–322.

[5] THORISSON, H. (1983). The coupling of regenerative processes. *Adv. in Appl. Probab.* **15** 531–561.

DEPARTMENT OF MATHEMATICS
CHALMERS UNIVERSITY OF TECHNOLOGY
AND THE UNIVERSITY OF GÖTEBORG
S-41296 GÖTEBORG, SWEDEN