The Annals of Probability
1989, Vol. 17, No. 1, 151–177

COUPLING METHODS FOR MULTIDIMENSIONAL DIFFUSION PROCESSES

BY MU-FA CHEN AND SHAO-FU LI

Beijing Normal University and Henan Teachers' University

In this paper, coupling methods for diffusion processes are studied mainly to obtain upper bound estimates in two different probability metrics. We use the martingale approach and explore the construction of explicit coupling operators which are sometimes optimal. The paper presents some criteria for the success of coupling and for the finiteness of the moments of the coupling times. Rates of convergence in various metrics are also studied.

1. Introduction. Coupling methods have been used widely in the study of interacting particle systems and other fields. There are some rather comprehensive treatments of coupling in the theory of Markov processes; see Griffeath (1978) and Liggett (1985). These papers contain a large number of references. In the diffusion context, we should mention Davies (1986), Lindvall (1983) and Lindvall and Rogers (1986). In the last paper, the authors obtained a successful coupling for a class of multidimensional diffusion processes by a reflection method and the theory of stochastic differential equations. Their method is effective for Brownian motion and for process in which the covariance matrix is almost constant. In particular, Brownian motion has a successful coupling. A geometric generalization of this Brownian coupling has been developed by Kendall (1986a, b) for use in stochastic differential geometry.

Let λ be a metric on $\mathbb{R}^d$. For $p \geq 1$, we define a probability metric W_p (often called Wasserstein or Kantorovich–Robinshtein–Wasserstein metric),

$$W_p(P_1, P_2) = \inf_Q \left[\int_{\mathbb{R}^d \times \mathbb{R}^d} \lambda(x, y)^p Q(dx, dy) \right]^{1/p},$$

where the infimum is taken over all measures Q on $\mathbb{R}^d \times \mathbb{R}^d$ such that for any measurable set $B \subset \mathbb{R}^d$,

$$\begin{aligned} Q(B \times \mathbb{R}^d) &= P_1(B), \\ Q(\mathbb{R}^d \times B) &= P_2(B). \end{aligned} \tag{1.1}$$

Any such Q is called a coupling of P_1 and P_2. Clearly, any coupling will give us an upper bound estimate for W_p. In this paper, we consider only the Euclidean metric on $\mathbb{R}^d$,

$$\rho(x, y) = |x - y| = \left[\sum_{i=1}^{d} (x_i - y_i)^2 \right]^{1/2}$$

Received July 1987; revised March 1988.
AMS 1980 *subject classifications.* Primary 60J60, 60J65, 60H10; secondary 60J45, 60J70.
Key words and phrases. Coupling, coupling operator, probability metric, coupling time, multidimensional diffusion, martingale approach.

and the discrete metric

$$d(x, y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

In the latter case, we will use $V(P_1, P_2)$ to distinguish W_1 from the metric ρ. Note that W_p is an analogue of the L^p-metric. It was proved by Dobrushin (1970) that

$$V(P_1, P_2) = \sup_B |P_1(B) - P_2(B)|,$$

which is just half of the total variation norm.

It will become clear later that different couplings are suitable for different metric. For this reason, we may use the terms "W_p-coupling" and "V-coupling," respectively, for the different purposes. The W_p-couplings are often used in the study of interacting particle systems. [See Chen (1986b, 1987b) and the references there.] More recently, the W_2-couplings have also been used in the study of infinite dimensional diffusion processes by J. M. Xu and the first author.

Now, let us consider the V-coupling. Suppose that $\{X_t\}_{t \geq 0}$ and $\{Y_t\}_{t \geq 0}$ are diffusion processes in $\mathbb{R}^d$ with the same transition function $P(t, \cdot, \cdot)$ and distributions P^x and P^y, respectively. Let $P^{x, y}$ be a coupling probability measure on $\Omega_{2d} = C([0, \infty); \mathbb{R}^{2d})$. That is, the first and the second d-dimensional (marginal) distributions of $P^{x, y}$ are the same as P^x and P^y, respectively. Define the coupling time as

$$T = \inf\{t \geq 0 \colon X(t) = Y(t)\}.$$

If

$$T < \infty, \quad P^{x, y}\text{-a.s.}$$

and

$$P^{x, y}[X(t) = Y(t);\ t \geq T] = 1,$$

we call the coupling $P^{x, y}$ successful. Furthermore, if

$$P^{x, y}[T > t] = o(t^{-\alpha}) \quad \text{as } t \to \infty,$$

for some $\alpha > 0$, then we have

$$V(P(t, x, \cdot), P(t, y, \cdot)) = o(t^{-\alpha}).$$

Thus the key point is to construct a successful coupling $P^{x, y}$ of P^x and P^y.

As we did in the case of jump processes [Chen (1986a, 1987a)], we begin our study with the analysis of coupling operators.

Let

$$L = \frac{1}{2} \sum_{i, j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial}{\partial x_i}$$

be an elliptic operator on $\mathbb{R}^d$ (possibly degenerate). Assume that the solution to the martingale problem for L is well-posed [see Stroock and Varadhan (1979)]. Our goal is to find an elliptic operator on $\mathbb{R}^{2d}$ such that the solution to the martingale problem for this operator has the marginal property (1.1). From the infinitesimal character of diffusion processes and (1.1), it is obvious that the coefficients of the operator should be of the form

$$a(x, y) = \begin{pmatrix} a(x) & c(x, y) \\ c(x, y)^* & a(y) \end{pmatrix}, \qquad b(x, y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix},$$

where c^* is the transpose of c. A trivial example is $c \equiv 0$. In this case, the two coordinates are independent and the coupling is usually not useful. But this means that a coupling operator always exists. Indeed, there are a lot of choices.

EXAMPLE 1.2 (d-dimensional Brownian motion). Take

$$\begin{aligned} c_1(x, y) &= I - 2(x - y)(x - y)^*/|x - y|^2, \\ c_2(x, y) &= I - (x - y)(x - y)^*/|x - y|^2, \\ c_3(x, y) &= \big((1 - \alpha(x_i - y_i)/(\beta + |x_i - y_i|))\delta_{ij}\big), \end{aligned}$$

where $\alpha \in (0, 2]$ and $\beta > 0$. All these couplings are successful (see Section 4). The first one was given in Lindvall and Rogers (1986), called coupling by reflection. Actually, if we denote by L_{xy} $(x \neq y)$ the hyperplane $\{z \in \mathbb{R}^d\colon (z, x - y) = 0\}$ which is just the orthogonal complement of $\{x - y\}$, then for each $z \in \mathbb{R}^d$, $c_1(x, y)z$ is the reflection image of z with respect to the hyperplane L_{xy}. On the other hand, $c_2(x, y)z$ is the projection of z onto the subspace L_{xy}. Hence we call the second one "coupling by projection." The last one has an advantage in that the couplings for different components are independent.

The paper is organized as follows: In the next section, we discuss the W_p-couplings ($p = 1, 2$). The remainder of the paper is devoted to the V-coupling which are much more complicated. In Section 3 we study the constructions of the couplings. In Section 4 we present some criteria for the success of couplings and a large number of examples to illustrate these criteria. Our criteria are exact in some cases. In Section 5 we study the rates of convergence of $P^{x,y}[T > t]$ as $t \to \infty$ for successful couplings. The moment of the coupling time T is also studied there.

2. Coupling for W_p-metric ($p = 1, 2$). Let $\Omega = \Omega_{2d} = C([0, \infty); \mathbb{R}^{2d})$ be the space of continuous trajectories from $[0, \infty)$ into $\mathbb{R}^{2d}$. Given $t \geq 0$ and $\omega \in \Omega$, let $Z(t, \omega) = Z_t(\omega)$ denote the position of ω in $\mathbb{R}^{2d}$. Define

$$\mathscr{M}_t = \sigma\{Z_s\colon s \leq t\}, \quad \mathscr{M} = \sigma\Big(\bigcup_{t \geq 0} \mathscr{M}_t\Big).$$

Let

$$\begin{aligned} X(t, \omega) &= \pi_1 \circ Z(t, \omega) = (\omega_1(t), \dots, \omega_d(t)), \\ Y(t, \omega) &= \pi_2 \circ Z(t, \omega) = (\omega_{d+1}(t), \dots, \omega_{2d}(t)). \end{aligned}$$

That is, $Z(t, \omega) = (X(t, \omega), Y(t, \omega))$. Similarly, we can define $\mathscr{M}_t^{(1)}, \mathscr{M}^{(1)}$ and

$\mathscr{M}_t^{(2)}$, $\mathscr{M}^{(2)}$. For example,

$$\mathscr{M}_t^{(1)} = \sigma\{X_s \colon s \le t\}.$$

We often denote the operator

$$L = \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2}{\partial x_i \, \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial}{\partial x_i}$$

by $L(a(x), b(x))$. Now suppose that $L_1(a_1(x), b_1(x))$ and $L_2(a_2(y), b_2(y))$ are given operators, then we can define an operator $L(a(x, y), b(x, y))$ on $\mathbb{R}^{2d}$ as

$$a(x, y) = \begin{pmatrix} a_1(x) & c(x, y) \\ c(x, y)^* & a_2(y) \end{pmatrix}, \qquad b(x, y) = \begin{pmatrix} b_1(x) \\ b_2(y) \end{pmatrix},$$

where $c(x, y)$ is a real valued $d \times d$ matrix such that the matrix $a(x, y)$ is nonnegative definite. Such an operator $L(a, b)$ is called a coupling of L_1 and L_2.

Throughout this paper, the coefficients of all operators are assumed to be locally bounded. Moreover, we assume that the martingale problem for the marginal diffusion processes are well posed. The solutions are denoted by

$$P_1^x \sim L_1, \qquad x \in \mathbb{R}^d;$$
$$P_2^y \sim L_2, \qquad y \in \mathbb{R}^d.$$

LEMMA 2.1. *Let $\{P^{x,y} \colon x, y \in \mathbb{R}^d\}$ be a family of solutions to the martingale problem for the coupling operator $L(a(x, y), b(x, y))$, denoted by $P^{x,y} \sim L(a, b)$. Then*

$$P_1^x = P^{x,y} \circ \pi_1^{-1}, \qquad P_2^y = P^{x,y} \circ \pi_2^{-1}, \qquad x, y \in \mathbb{R}^d.$$

In other words, $P^{x,y}$ is a coupling of P_1^x and P_2^y for every $x, y \in \mathbb{R}^d$.

PROOF. Let $f \in C_0^\infty(\mathbb{R}^d)$ and set $F(x, y) = f(x)$, $x, y \in \mathbb{R}^d$. Note that $LF(x, y) = L_1 f(x)$ and the operators are locally bounded. We have, for every set $B \in \mathscr{M}_s^{(1)}$ and $s \le t$, that $\pi_1^{-1}B \in \mathscr{M}_s$ and

$$\begin{aligned}
&\int_B \left(f(X_t) - \int_0^t L_1 f(X_u) \, du \right) d\left(P^{x,y} \circ \pi_1^{-1}\right) \\
&= \int_{\pi_1^{-1}B} \left(F(Z_t) - \int_0^t LF(Z_u) \, du \right) dP^{x,y} \\
&= \int_{\pi_1^{-1}B} \left(F(Z_s) - \int_0^s LF(Z_u) \, du \right) dP^{x,y} \\
&= \int_B \left(f(X_s) - \int_0^s L_1 f(X_u) \, du \right) d\left(P^{x,y} \circ \pi_1^{-1}\right).
\end{aligned}$$

This shows that $P^{x,y} \circ \pi_1^{-1} \sim L_1$. By the uniqueness assumption, we certainly get

$$P_1^x = P^{x,y} \circ \pi_1^{-1}, \qquad x, y \in \mathbb{R}^d.$$

Similarly, we have the second equality. □

We need the following elementary result.

LEMMA 2.2. *Let $V(t)$ be a differentiable function and $B(t)$ be a locally integrable function on $[0, \infty)$. If*

$$\frac{d}{dt}V(t) \le -cV(t) + B(t), \quad a.e.\ t,$$

for some $c > 0$, then

$$V(t) \le V(0)e^{-ct} + \int_0^t e^{-c(t-s)}B(s)\,ds, \qquad t \ge 0.$$

THEOREM 2.3. *Suppose that $a(x, y)$ and $b(x, y)$ are continuous on $\mathbb{R}^{2d}$ and $P^{x,y} \sim L(a, b)$. If there exist constants $C \ge 0$ and $c > 0$ such that*

$$L\rho^2(x, y) \le C - c\rho^2(x, y), \qquad x, y \in \mathbb{R}^d,$$

then

$$E^{x,y}\rho^2(X_t, Y_t) \le C/c + e^{-ct}\rho^2(x, y), \qquad x, y \in \mathbb{R}^d.$$

In particular, if $C = 0$, then

$$W_2(P_1(t, x, \cdot), P_2(t, y, \cdot)) \le \rho(x, y)e^{-ct/2} \to 0 \quad \text{as } t \to \infty,$$

where $P_1(t, x, \cdot)$ and $P_2(t, y, \cdot)$ are, respectively, the transition functions of the marginal diffusions. The same conclusion is true if we replace ρ^2, W_2 and $e^{-ct/2}$ by ρ, W_1 and e^{-ct}, respectively.

PROOF. Set

$$S_N = \inf\{t \ge 0: |X_t - Y_t| > N\},$$
$$T_R = \inf\{t \ge 0: |X_t|^2 + |Y_t|^2 > R\},$$
$$S = S_N \wedge T_R.$$

Since $P^{x,y} \sim L$, we have

$$E^{x,y}\rho^2(X_{t\wedge S}, Y_{t\wedge S}) = \rho^2(x, y) + \int_0^t E^{x,y}L\rho^2(X_{u\wedge S}, Y_{u\wedge S})\,du$$

and so

$$\frac{d}{dt}E^{x,y}\rho^2(X_{t\wedge S}, Y_{t\wedge S}) = E^{x,y}L\rho^2(X_{t\wedge S}, Y_{t\wedge S})$$
$$\le C - cE^{x,y}\rho^2(X_{t\wedge S}, Y_{t\wedge S}).$$

By Lemma 2.2 we obtain

$$E^{x,y}\rho^2(X_{t\wedge S}, Y_{t\wedge S}) \le C/c + \rho^2(x, y)e^{-ct}.$$

Now the conclusion follows by passing the limit $R \uparrow \infty$, $N \uparrow \infty$. □

DEFINITION 2.4. Let $a_1(x) = \sigma_1(x)\sigma_1(x)^*$, $a_2(y) = \sigma_2(y)\sigma_2(y)^*$. We call

$$a(x, y) = \begin{pmatrix} a_1(x) & \sigma_1(x)\sigma_2(y)^* \\ \sigma_2(y)\sigma_1(x)^* & a_2(y) \end{pmatrix}, \qquad b(x, y) = \begin{pmatrix} b_1(x) \\ b_2(y) \end{pmatrix}$$

the basic coupling of L_1 and L_2.

EXAMPLE 2.5 [Ornstein–Uhlenbeck (O.U.) process].

$$\sigma_1(x) = \sigma_2(x) = I, \qquad b_1(x) = b_2(x) = -\mu x.$$

Using the basic coupling, we obtain from Theorem 2.3

$$W_2(P(t, x, \cdot), P(t, y, \cdot)) \le \left(E^{x, y}\rho^2(X_t, Y_t)\right)^{1/2} = e^{-\mu t}\rho(x, y),$$

$$W_1(P(t, x, \cdot), P(t, y, \cdot)) \le E^{x, y}\rho(X_t, Y_t) = e^{-\mu t}\rho(x, y), \qquad t \ge 0,\ x, y \in \mathbb{R}^d.$$

EXAMPLE 2.6. Take $\sigma_1(x) = \sigma_2(x) = \sigma =$ constant, $b_1(x) = b_2(x) = 0$. Using the basic coupling, we obtain from Theorem 2.3

$$W_2(P(t, x, \cdot), P(t, y, \cdot)) \le \left(E^{x, y}\rho^2(X_t, Y_t)\right)^{1/2} = |x - y|,$$

$$W_1(P(t, x, \cdot), P(t, y, \cdot)) \le E^{x, y}\rho(X_t, Y_t) = |x - y|.$$

On the other hand, it is known from Givens and Shortt (1984) that the first inequality is an equality in any dimension. Thus, our basic coupling is exact in this case. For W_1, the coupling is not exact, but as you will see soon it is the best that we can do.

We now introduce some notation which will be used often later.

2.7 NOTATION. Denote by $\langle\, ,\rangle$ the ordinary inner product in $\mathbb{R}^d$. Set

$$A(x, y) = a_1(x) + a_2(y) - 2c(x, y),$$
$$B(x, y) = b_1(x) - b_2(y),$$
$$\hat{A}(x, y) = \langle x - y, A(x, y)(x - y)\rangle,$$
$$\bar{A}(x, y) = \hat{A}(x, y)/|x - y|^2, \qquad x \ne y,$$
$$\hat{B}(x, y) = \langle x - y, B(x, y)\rangle.$$

It is easy to check that $\hat{A}(x, y) \ge 0$ for all $x, y \in \mathbb{R}^d$ [since $a(x, y)$ is nonnegative definite] and that for each $f \in C^2([0, \infty))$, we have

$$\begin{aligned} 2Lf(\rho(x, y)) = {} & \bar{A}(x, y)f''(\rho(x, y)) + f'(\rho(x, y)) \\ & \times\left[\operatorname{tr} A(x, y) - \bar{A}(x, y) + 2\hat{B}(x, y)\right]/\rho(x, y). \end{aligned} \tag{2.8}$$

In particular, we have

$$L\rho^2(x, y) = \operatorname{tr} A(x, y) + 2\hat{B}(x, y) \tag{2.9}$$

and

$$(2.10)\qquad L\rho(x, y) = \left[\operatorname{tr} A(x, y) - \overline{A}(x, y) + 2\hat{B}(x, y)\right]/(2\rho(x, y)).$$

Now, we turn to discuss how to choose the coupling operators for $W_2(W_1)$-coupling. For simplicity, we consider only the case that $b_1(x) = b_2(y) = 0$.

REMARK 2.11. In view of Theorem 2.3, (2.9) and (2.10), we may say that a coupling operator $a(x, y)$ is W_2- (respectively, W_1-) optimal if $a(x, y)$ is nonnegative definite and $\operatorname{tr} A(x, y)$ [respectively, $\operatorname{tr} A(x, y) - \overline{A}(x, y)$] achieves the minimum at each point $(x, y) \in \mathbb{R}^{2d}$ [note that these quantities contain $c(x, y)$ which varies]. Clearly, if $\sigma_1(x) = \sigma_2(y) = \sigma =$ constant, then the basic coupling gives us $\operatorname{tr} A = \overline{A} = 0$ and so is optimal. For the general case, let us fix x and y, assume that $a_1(x)$ and $a_2(y)$ are positive definite and take $\sigma_1(x) = \sqrt{a_1(x)}$, $\sigma_2(y) = \sqrt{a_2(y)}$, the positive definite square roots. In this case, we can rewrite $c(x, y)$ as $\sigma_1(x)H^*(x, y)\sigma_2(y)$. Now, $a(x, y)$ is nonnegative definite if and only if H is contractive. That is, $|Hx| \le |x|$ for all $x \in \mathbb{R}^d$. Using the Hilbert–Schmidt (H.S.) norm for metrics, we can easily prove that the optimal choice of $H(x, y)$ does exist since the domain of H is compact and $\operatorname{tr} A$ (respectively, $\operatorname{tr} A - \overline{A}$) is continuous in H with respect to the H.S. norm. But this optimalization problem is generally quite difficult. For simplicity, now we restrict ourselves to the case that H is an orthogonal matrix. Then, for W_2, the solution is

$$(2.12)\qquad H(x, y) = \left[\sigma_2(y)a_1(x)\sigma_2(y)\right]^{-1/2}\sigma_2(y)\sigma_1(x).$$

Then, we have

$$\operatorname{tr} A(x, y) = \operatorname{tr}\left[a_1(x) + a_2(y) - 2\left(\sigma_2(y)a_1(x)\sigma_2(y)\right)^{1/2}\right].$$

(In this case, even without the orthogonal assumption on H, the optimal solution is still the same. For details, see Givens and Shortt [(1984), pages 237–239].) Furthermore, if $\sigma_1 = \sigma_2 = \sigma$ is diagonal, then we have $H(x, y) = (\sigma_{ii}(x)\sigma_{ii}(y)\delta_{ij})$. For W_1, the optimal solution H should satisfy

$$(2.13)\qquad H\sigma_1(I - \bar{u}\bar{u}^*)\sigma_2 = \left[\sigma_2(I - \bar{u}\bar{u}^*)a_1(I - \bar{u}\bar{u}^*)\sigma_2\right]^{1/2},$$

where $\bar{u} = (x - y)/|x - y|$, $H = H(x, y)$, $\sigma_1 = \sigma_1(x)$, $\sigma_2 = \sigma_2(y)$ and so on. This is quite complicated but still useful in some cases. We will return to this formula later.

A typical application of the above coupling is as follows: Suppose that our diffusion with L_1 has a stationary distribution π, the conditions of Theorem 2.3 are satisfied with $C = 0$ for a coupling of L_1 and itself, and $(x, y) \to P^{x, y}$ is measurable, then we have

$$W_2(P(t, x, \cdot), \pi) \le e^{-ct/2}\int \pi(dy)|x - y|.$$

In fact,

$$
\begin{aligned}
W_2(P(t,x,\cdot),\pi) &= W_2\Big(P(t,x,\cdot),\int\pi(dy)P(t,y,\cdot)\Big)\\
&\le \Big[\int\pi(dy)E^{x,y}|X_t-Y_t|^2\Big]^{1/2} \qquad (2.14)\\
&\le e^{-ct/2}\Big[\int\pi(dy)|x-y|^2\Big]^{1/2},
\end{aligned}
$$

and similarly for W_1. As for the existence of stationary distribution for diffusions, see Bhattacharya and Ramasubramanian (1982) and their references. Here, we state a simple sufficient condition. The proof is nontrivial but almost the same as the ones in Basis (1980) and Chen (1986b) which go back to Dobrushin (1970). Hence we omit the proof.

THEOREM 2.15. *Let $h \in C^2(\mathbb{R}^d)$ be a compact function, i.e., $h \ge 0$, $\{x\colon h(x) \le k\}$ is a compact set for each $k \ge 0$. If there are constants $C \ge 0$ and $c > 0$ such that*

$$L_1 h(x) \le C - ch(x), \qquad x \in \mathbb{R}^d, \tag{2.16}$$

then the diffusion process determined by L_1 has a stationary distribution π with

$$\int\pi(dx)h(x) \le C/(1-c). \tag{2.17}$$

Now, if (2.16) holds with $h = \rho^2$, then combining (2.14) and (2.17), we obtain

$$W_2(P(t,x,\cdot),\pi) \le \text{const.}\,(1+|x|)e^{-ct/2} \to 0 \quad \text{as } t \to \infty. \tag{2.18}$$

3. Constructions of couplings for V-metric. Starting with this section, we discuss the couplings for V-metric. The following result describes a fundamental property of our basic coupling.

THEOREM 3.1. *Let $a_1 = a_2 = \sigma\sigma^*$, $b_1 = b_2 = b$, and σ and b be continuous on $\mathbb{R}^d$. Suppose that for the basic coupling $L(a(x,y), b(x,y))$:*

$$a(x,y) = \begin{pmatrix} \sigma(x)\sigma(x)^* & \sigma(x)\sigma(y)^* \\ \sigma(y)\sigma(x)^* & \sigma(y)\sigma(y)^* \end{pmatrix}, \qquad b(x,y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix},$$

the martingale problem is locally well-posed [and hence globally well-posed by Stroock and Varadhan (1979), Corollary 10.1.2]. If we denote the solution by $P^{x,y} \sim L(a(x,y), b(x,y))$, then we have

$$X_t = Y_t, \qquad t \ge T, \qquad P^{x,y}\text{-}a.s. \text{ on } [T<\infty]. \tag{3.2}$$

PROOF. A similar result was given in Stroock and Varadhan [(1979), Lemma 8.1.3]. Here we present a different proof. By a modification of Theorem 6.1.3 in

Stroock and Varadhan (1979) (we allow $T = \infty$), what we need is to show that

$$P^{x,x}[X_t = Y_t,\ t \ge 0] = 1, \qquad x \in \mathbb{R}^d.$$

Next, by a truncation argument, we may assume that $a(x, y)$ and $b(x, y)$ are bounded and continuous, so the martingale problem for L is well-posed.

Now take

$$\eta(x) = \begin{cases} C\exp[-1/(1-|x|^2)], & |x| < 1,\ x \in \mathbb{R}^d, \\ 0, & |x| \ge 1,\ x \in \mathbb{R}^d, \end{cases}$$

where C is the normalizing constant. Put $\eta_\varepsilon(x) = \varepsilon^{-d}\eta(x/\varepsilon)$, set $\sigma_\varepsilon(x) = (\sigma_{ij}^\varepsilon(x))$: $\sigma_{ij}^\varepsilon(x) = (\sigma_{ij} * \eta_\varepsilon)(x)$, $b_\varepsilon(x) = (b_i^\varepsilon(x))$: $b_i^\varepsilon(x) = (b_i * \eta_\varepsilon)(x)$ and so on. Clearly $\sigma_{ij}^\varepsilon(x), b_i^\varepsilon(x) \in C_b^\infty(\mathbb{R}^d)$. Hence there exist constants A_ε and B_ε such that

$$\|\sigma_\varepsilon(x) - \sigma_\varepsilon(y)\| \le A_\varepsilon|x - y|,$$
$$|b_\varepsilon(x) - b_\varepsilon(y)| \le B_\varepsilon|x - y|.$$

On the other hand, corresponding to the function $\rho(x, y) = |x - y|$, we can construct a sequence of functions $\{\phi_n\}_1^\infty$ such that

$$\phi_n \in C^2(\mathbb{R}), \qquad \phi_n(x) \uparrow |x|, \qquad |\phi_n'| \le 1, \qquad 0 \le \phi_n''(x) \le 2/(nx^2).$$

[See Ikeda and Watanabe (1981), pages 168–169.] Now, let $L_\varepsilon(\sigma_\varepsilon^*, b_\varepsilon)$ be the basic coupling. Then by (2.8), we have

$$2L_\varepsilon\phi_n(|x-y|) = \phi_n''(|x-y|)\bar{A}_\varepsilon(x, y) + \frac{\phi_n'(|x-y|)}{|x-y|}\big(\operatorname{tr} A_\varepsilon(x, y) - \bar{A}_\varepsilon(x, y) + 2\hat{B}_\varepsilon(x, y)\big),$$

where

$$\bar{A}_\varepsilon(x, y) = \left|(\sigma_\varepsilon(x) - \sigma_\varepsilon(y))^* \frac{x - y}{|x - y|}\right|^2 \le A_\varepsilon^2|x - y|^2,$$
$$\operatorname{tr} A_\varepsilon(x, y) \le \|\sigma_\varepsilon(x) - \sigma_\varepsilon(y)\|^2 \le A_\varepsilon^2|x - y|^2,$$
$$|\hat{B}_\varepsilon(x, y)| = |\langle x - y, b_\varepsilon(x) - b_\varepsilon(y)\rangle| \le B_\varepsilon|x - y|^2$$

and so

$$2L_\varepsilon\phi_n(|x - y|) \le A_\varepsilon^2\phi_n''(|x-y|)|x-y|^2 + |\phi_n'(|x-y|)|(2A_\varepsilon^2 + 2B_\varepsilon)|x - y|.$$

Let $P_\varepsilon^{x,x} \sim L_\varepsilon(\sigma_\varepsilon\sigma_\varepsilon^*, b_\varepsilon)$ and

$$S_N = \inf\{t \ge 0: |X_t - Y_t| > N\}.$$

Then

$$E_\varepsilon^{x,x}\phi_n(|X_{t\wedge S_N} - Y_{t\wedge S_N}|) \le E_\varepsilon^{x,x}\int_0^{t\wedge S_N}\{\tfrac{1}{2}A_\varepsilon^2\phi_n''\rho^2 + |\phi_n'|(A_\varepsilon^2 + B_\varepsilon)\rho\}(X_u, Y_u)\,du$$
$$\le A_\varepsilon^2 t/n + (A_\varepsilon^2 + B_\varepsilon)E_\varepsilon^{x,x}\int_0^{t\wedge S_N}|X_u - Y_u|\,du.$$

Let $N \uparrow \infty$ and then $n \uparrow \infty$. We obtain

$$E_\varepsilon^{x,x}|X_t - Y_t| \le \left(A_\varepsilon^2 + B_\varepsilon\right)\int_0^t E_\varepsilon^{x,x}|X_u - Y_u|\,du$$

and so

$$E_\varepsilon^{x,x}|X_t - Y_t| = 0, \qquad t \ge 0.$$

This shows that

$$P_\varepsilon^{x,x}(X_t = Y_t,\ t \ge 0) = 1.$$

Finally, by Stroock and Varadhan (1979), Theorem 1.4.6, it is easy to prove that $\{P_\varepsilon^{x,y}\colon \varepsilon > 0\}$ is tight, and so we can choose a subsequence $\{\varepsilon_m\}_1^\infty$ such that

$$P_{\varepsilon_m}^{x,y} \to P_0^{x,y} \quad \text{weakly}.$$

Then $P_0^{x,y} \sim L(a(x,y), b(x,y))$. By the locally well-posed assumption, we indeed have $P_0^{x,y} = P^{x,y}$. Therefore

$$P^{x,x}[X_t = Y_t] \ge \limsup_{m\to\infty} P_{\varepsilon_m}^{x,x}[X_t = Y_t] = 1, \qquad t \ge 0.$$

This proves our assertion. □

In order to give different types of couplings, we need more preparation. Denote by $\tilde{P}^{x,y}$ the solution to the martingale problem for the basic coupling constructed by Theorem 3.1. Set

$$Q_\omega = \delta_\omega \bigotimes\nolimits_{T(\omega)} \tilde{P}^{X(T(\omega)),\,Y(T(\omega))} I_{[T(\omega)<\infty]} + \delta_\omega I_{[T(\omega)=\infty]}, \qquad \omega \in \Omega. \tag{3.3}$$

Lemma 3.4. *Under the hypotheses of Theorem* 3.1, *if* $P^{x,y}$ *is a solution to the martingale problem for*

$$a(x,y) = \begin{pmatrix} \sigma(x)\sigma(x)^* & c(x,y) \\ c(x,y)^* & \sigma(y)\sigma(y)^* \end{pmatrix}, \qquad b(x,y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix}$$

up to time T, *then*

$$R = P^{x,y} \bigotimes\nolimits_T Q$$

is a solution to the martingale problem for

$$a(t,x,y) = \begin{pmatrix} \sigma(x)\sigma(x)^* & \begin{array}{l} I_{[0,T)}(t)c(x,y) \\ \quad + I_{[T,\infty)}(t)\sigma(x)\sigma(y)^* \end{array} \\ \begin{array}{l} I_{[0,T)}(t)c(x,y)^* \\ \quad + I_{[T,\infty)}(t)\sigma(y)\sigma(x)^* \end{array} & \sigma(y)\sigma(y)^* \end{pmatrix}, \tag{3.5}$$

$$b(t,x,y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix}.$$

Proof. Cf. Stroock and Varadhan [(1979), Section 6.1] for details. □

EXAMPLE 3.6 (Classical coupling). In (3.5), take $c(x, y) = 0$. This means that the processes start from two different points, run independently until they first meet each other, then move together.

EXAMPLE 3.7 (Coupling by reflection). Take

$$c(x, y) = \sigma(x)(I - 2\bar{u}\bar{u}^*)\sigma(y)^*,$$

where $\bar{u} = (x - y)/|x - y|$. If σ is constant and $\det \sigma \neq 0$, we can also take

$$c(x, y) = \sigma\sigma^* - 2\bar{u}\bar{u}^*/|\sigma^{-1}\bar{u}|^2.$$

EXAMPLE 3.8 (Coupling by projection. Take

$$c(x, y) = \sigma(x)(I - \bar{u}\bar{u}^*)\sigma(y)^*.$$

For the above examples, we can first construct the couplings up to time T, then applying Lemma 3.4, link them with the basic coupling so that after time T, they will move together. Sometimes, we have to do so (cf. Section 4). However, it is not always the case. Very often, it is enough to construct a coupling up to the time T. This also enables us to consider the more general case that $L_1 \neq L_2$. Such generalization is useful in some cases [see Chen (1986b), for example].

4. Criteria for success. In this and the next sections, we fix a coupling operator $L(a(x, y), b(x, y))$:

$$a(x, y) = \begin{bmatrix} a_1(x) & c(x, y) \\ c(x, y)^* & a_2(y) \end{bmatrix}, \qquad b(x, y) = \begin{bmatrix} b_1(x) \\ b_2(y) \end{bmatrix},$$

and assume that $P^{x,y}$ $(x \neq y)$ is a solution to the martingale problem for L up to time T:

$$T = \inf\{t \geq 0: X_t = Y_t\}.$$

In other words, for each pair $x \neq y$ and for every $f \in C_0^2(\mathbb{R}^{2d})$: $\operatorname{supp}(f) \subseteq \{(x, y) \in \mathbb{R}^{2d}: 1/n \leq |x - y| \leq N\}$ for some $n, N > 1$,

$$f(X_t, Y_t) - \int_0^t Lf(X_u, Y_u)\, du$$

is a $P^{x,y}$-martingale with respect to $\{\mathscr{M}_t\}_{t \geq 0}$.

The idea of our criteria discussed below is to compare the process $Z_t = (X_t, Y_t)$ with the radial process $r_t = |X_t - Y_t|$. To do this, set

$$S_N = \inf\{t \geq 0: |X_t - Y_t| > N\}, \qquad N > 1,$$
$$T_n = \inf\{t \geq 0: |X_t - Y_t| < 1/n\}, \qquad n > 1,\ T_n \uparrow T \text{ as } n \uparrow \infty,$$
$$T_{n,N} = T_n \wedge S_N.$$

Choose continuous functions γ and γ^*: $(0, \infty) \to \mathbb{R}$ such that

$$\gamma(r) \geq \sup_{|x-y|=r} \left(\operatorname{tr} A(x, y) - \bar{A}(x, y) + 2\hat{B}(x, y)\right)/\bar{A}(x, y),$$

$$\gamma_*(r) \leq \inf_{|x-y|=r} \left(\operatorname{tr} A(x, y) - \bar{A}(x, y) + 2\hat{B}(x, y)\right)/\bar{A}(x, y)$$

and define

$$C(r) = \exp\left[\int_1^r \frac{\gamma(u)}{u}\,du\right], \qquad C_*(r) = \exp\left[\int_1^r \frac{\gamma_*(u)}{u}\,du\right],$$

$$f(r) = \int_1^r C(s)^{-1}\,ds, \qquad f_*(r) = \int_1^r C_*(s)^{-1}\,ds, \quad r > 0.$$

Then, we have

$$\begin{aligned} f'(r) &> 0, \qquad f''(r) + f'(r)\gamma(r)/r = 0, \\ f_*'(r) &> 0, \qquad f_*''(r) + f_*'(r)\gamma_*(r)/r = 0. \end{aligned} \tag{4.1}$$

Next, choose continuous functions α and α^*: $(0, \infty) \to [0, \infty)$ such that

$$\alpha(r) \le \inf_{|x-y|=r} \overline{A}(x, y) \le \sup_{|x-y|=r} \overline{A}(x, y) \le \alpha^*(r).$$

Define

$$g(r) = \int_r^1 C(s)^{-1}\,ds \int_s^1 \frac{C(u)}{\alpha(u)}\,du,$$

$$g_*(r) = \int_r^1 C_*(s)^{-1}\,ds \int_s^1 \frac{C_*(u)}{\alpha^*(u)}\,du,$$

as $r \uparrow \infty$, $f(r) \uparrow f(\infty)$, say. Similarly, we can define $f(0)$, $f_*(\infty)$, $f_*(0)$, $g(0)$ and $g_*(0)$.

THEOREM 4.2. *Let $\alpha > 0$ on $(0, \infty)$.*

(i) *If $f(\infty) = \infty$ and $g(0) < \infty$, then the coupling is successful.*

(ii) *If $f_*(\infty) < \infty$ or $g_*(0) = \infty$, then the coupling is not successful.*

(iii) *If $\gamma = \gamma_*$ and $\alpha = \alpha^*$, then the coupling is successful if and only if $f(\infty) = \infty$ and $g(0) < \infty$.*

COROLLARY 4.3. (i) *If α is bounded below by a positive number, $f(\infty) = \infty$, $f(0) > -\infty$ and $\liminf_{\gamma \downarrow 0} f'(r) > 0$, then the coupling is successful.*

(ii) *If $\alpha > 0$ on $(0, \infty)$, $f_*(\infty) < \infty$ or $f_*(0) = -\infty$, then the coupling is not successful.*

PROOF. In case (i), it is easy to check that $g(0) < \infty$. As for case (ii), it suffices to note that $f_*(0) = -\infty \Rightarrow g_*(0) = \infty$. Thus, the corollary follows from Theorem 4.2 directly. □

Case (i) of Corollary 4.3 was obtained by Lindvall and Rogers [(1986), Lemma 1].

PROOF OF THEOREM 4.2. For the sake of completeness and also for certain subsequent uses, we sketch the proof here though the technique is essentially not new [cf. Friedman (1975)].

Set

$$(4.4)\quad F_{n,N}(\rho) = -\int_{1/n}^{\rho} C(s)^{-1}\int_s^N \frac{C(u)}{\alpha(u)}\,du, \qquad \frac{1}{n} \le \rho \le N,\ n, N > 1.$$

Then

$$(4.5)\quad \begin{gathered} -\infty < F_{n,N}(\rho) \le 0, \qquad F'_{n,N}(\rho) \le 0, \\ F''_{n,N}(\rho) + F'_{n,N}(\rho)\gamma(\rho)/\rho = 1/\alpha(\rho). \end{gathered}$$

Combining this with (2.8), we have

$$(4.6)\quad 2LF_{n,N}(\rho(x,y)) \ge 1.$$

Put $r = |x - y|$. Since $P^{x,y} \sim L(a, b)$, by a truncation argument we have

$$\begin{aligned} E^{x,y}F_{n,N}\big(|X_{t\wedge T_{n,N}} - Y_{t\wedge T_{n,N}}|\big) - F_{n,N}(r) &= \tfrac{1}{2}E^{x,y}\int_0^{t\wedge T_{n,N}} 2LF(|X_u - Y_u|)\,du \\ &\ge \tfrac{1}{2}E^{x,y}(t\wedge T_{n,N}), \end{aligned}$$

and so

$$E^{x,y}(t\wedge T_{n,N}) \le -2F_{n,N}(r).$$

Letting $t \uparrow \infty$, we get

$$(4.7)\quad E^{x,y}(T_{n,N}) \le -2F_{n,N}(r) < \infty.$$

(i) If $g(0) < \infty$, then

$$F_{0,N}(r) \equiv \lim_{n\to\infty} F_{n,N}(r) > -\infty.$$

From (4.7), it follows that

$$(4.8)\quad E^{x,y}(T \wedge S_N) \le -2F_{0,N}(r) < \infty.$$

On the other hand, since $Lf(\rho(x, y)) \le 0$, we have

$$(4.9)\quad \begin{gathered} f\left(\frac{1}{n}\right)P^{x,y}(T_n < S_n \wedge t) + f(N)P^{x,y}(S_N < T_n \wedge t) \\ + E^{x,y}\big(f(\rho(x,y)) : t \le T_{n,N}\big) \le f(r). \end{gathered}$$

Hence, by (4.7) we get

$$f\left(\frac{1}{n}\right)P^{x,y}(T_n < S_N) + f(N)P^{x,y}(S_N < T_n) \le f(r).$$

Thus

$$P^{x,y}(T_n > S_N) \le \frac{f(r) - f(1/n)}{f(N) - f(1/n)}.$$

Noting that $g(0) < \infty \Rightarrow f(0) > -\infty$, we have

$$P^{x,y}(T > S_N) \le \frac{f(r) - f(0)}{f(N) - f(0)}.$$

Letting $N \uparrow \infty$ and using (4.8), we obtain

$$P^{x,y}(T = \infty) = 0.$$

(ii) First, we assume that $f_*(\infty) < \infty$. By (4.7), we have

$$f_*\left(\frac{1}{n}\right)P^{x,y}(T_n < S_N) + f_*(N)P^{x,y}(T_n > S_N) \le f_*(r).$$

Hence

$$P^{x,y}(T_n < S_N) \le \frac{f_*(N) - f_*(r)}{f_*(N) - f_*(1/n)}$$

and so

$$P^{x,y}[T < \infty] \le P^{x,y}(T_n < \infty) \le \frac{f_*(\infty) - f_*(r)}{f_*(\infty) - f_*(1/n)} < 1.$$

Next, we assume that $g_*(0) = \infty$. Set

$$g_*(\rho, N) = \int_\rho^N C_*(s)^{-1}\,ds \int_s^N \frac{C(u)}{\alpha^*(u)}\,du < \infty, \qquad 0 < \rho \le N.$$

For $0 < \rho \le N$, set $g_*^{(0)}(\rho, N) = 1$ and define

$$g_*^{(m)}(\rho, N) = \int_\rho^N C_*(u)^{-1}\,ds \int_s^N \frac{C_*(u)}{\alpha^*(u)} g_*^{(m)}(u, N)\,du$$

inductively. Then, it is easy to check that

$$g_*^{(m)}(\rho, N) \le \frac{1}{m!} g_*(\rho, N)^m, \qquad m \ge 0.$$

Hence

$$u_N(\rho) = \sum_{m=0}^{\infty} g_*^{(m)}(\rho, N)$$

is well-defined for all $\rho \in (0, N]$. Moreover,

$$u_N \ge 1, \qquad u_N' \le 0,$$
$$1 + g_*(\rho, N) \le u_N(\rho) \le \exp(g_*(\rho, N)),$$
$$u_N(\rho) = \alpha^*(\rho)\left(-u_N''(\rho) + \frac{\gamma_*(r)}{r} u_N'(\rho)\right)$$

and so

$$\lim_{\rho \downarrow 0} u_N(\rho) = \infty, \qquad 2Lu_N(\rho(x, y)) \le u_N(\rho(x, y)).$$

Finally, fix $x \neq y$ and set $|x - y| = r > 0$. Then, by a truncating argument, for every $N > r$, we have

$$u_N(r) \geq E^{x,y}\left[e^{-T_{n,N} \wedge t/2} u_N\left(\rho\left(x_{T_{n,N} \wedge t}, Y_{T_{n,N} \wedge t}\right)\right)\right]$$

$$\geq E^{x,y}\left[e^{-t/2} u_N\left(\rho\left(X_{T_{n,N} \wedge t}, Y_{T_{n,N} \wedge t}\right)\right) : T_n \leq S_N \wedge t\right]$$

$$= u_N\left(\frac{1}{n}\right) e^{-t/2} P^{x,y}(T_n \leq S_N \wedge t),$$

that is,

$$P^{x,y}(T_n \leq S_N \wedge t) \leq e^{t/2} u_N(r) / u_N\left(\frac{1}{n}\right).$$

Letting $n \to \infty$ and then $N \to \infty$, we get

$$P^{x,y}(T \leq t) = 0, \qquad t \geq 0.$$

This gives us

$$P^{x,y}(T < \infty) = 0.$$

Equivalently,

$$P^{x,y}(T = \infty) = 1. \qquad \square$$

Example 4.10 [Classical coupling of Brownian motion (B.M.) in $\mathbb{R}^d$]. We have $\gamma(r) = \gamma_*(r) = d - 1$, $\alpha(r) = \alpha^*(r) = 2$. Hence

$$f(r) = \begin{cases} r - 1, & d = 1, \\ \log r, & d = 2, \\ (1 - r^{-d+2})/(d - 2), & d \geq 3, \end{cases}$$

$$g(r) = \begin{cases} \frac{1}{4}(1 - r)^2, & d = 1, \\ \frac{1}{4}\left(-\log r - \frac{1}{2} + \frac{r^2}{2}\right), & d = 2, \end{cases}$$

and so the coupling is successful if and only if $d = 1$. This result should come as no surprise since Brownian motion does not hit points in $d \geq 2$.

Example 4.11 (Coupling of B.M. in $\mathbb{R}^d$ by reflection or projection). In both cases, we have $\gamma \equiv 0$ and α = positive constant. Thus, $f(r) = r - 1$, $f'(r) = 1 > 0$ and so these couplings are successful.

Example 4.12 (Coupling of different diffusions). Take $d = 1$, $a_1(x) = a_1 > 0$, $a_2(y) = a_2 > 0$, $b_1(x) = -b_1 x$, $b_2(y) = -b_1 y - b_2$, $b_1, b_2 > 0$. Using the coupling by reflection, we get

$$\alpha(u) \equiv \alpha = \left(\sqrt{a_1} + \sqrt{a_2}\right)^2,$$

$$\gamma(u) = \frac{2}{\alpha}\left(-b_1 u^2 + b_2 u\right),$$

$$f(r) = \exp\left(-\frac{(b_1 - b_2)^2}{\alpha b_1}\right)\int_1^r \exp\left[\frac{b_1}{\alpha}\left(u - \frac{b_2}{b_1}\right)^2\right] du,$$

$$f'(r) > 0.$$

Hence the coupling is successful. If $a_1 \neq a_2$, then the basic coupling is also successful.

REMARK 4.13. Based on the idea of reflection, Lindvall and Rogers (1986) proposed a coupling by taking

$$c(x, y) = \sigma(x)\left(\sigma(y)^* - 2\frac{\sigma(y)^{-1}(x - y)(x - y)^*}{|\sigma^{-1}(y)(x - y)|^2}\right).$$

Under some hypotheses, they proved that this coupling satisfies the conditions of (i) of Corollary 4.3, so is successful. Since the hypotheses of Theorem 4.2 for success are weaker than those given in Corollary 4.3, our criterion is applicable to their case.

EXAMPLE 4.14. Take $\sigma(x) = \sqrt{2ax}$, $b(x) = cx + d$, $x \geq 0$, $a > 0$ and $d \geq 0$. The diffusion process on $[0, \infty)$ for this operator is well-defined [cf. Ikeda and Watanabe (1981), pages 221–222]. Use the coupling by reflection,

$$c(x, y) = -2a\sqrt{xy}.$$

We have

$$A(x, y) = \operatorname{tr} A(x, y) = \bar{A}(x, y) = 2a\left(\sqrt{x} + \sqrt{y}\right)^2,$$

$$\begin{aligned}\alpha(r) &= 2a \inf_{|x-y|=r} \left(\sqrt{x} + \sqrt{y}\right)^2 \\ &= 2a \inf_{x \geq 0} \left(\sqrt{x} + \sqrt{x + r}\right)^2 \\ &= 2ar,\end{aligned}$$

$$\gamma(u) = \frac{c}{a}u,$$

$$C(s) = \exp\left[\frac{c}{a}(s - 1)\right],$$

$$f(r) = \begin{cases} r - 1, & c = 0, \\ \dfrac{a}{c}\left[1 - \exp\left[\dfrac{c}{a}(1 - r)\right]\right], & c \neq 0. \end{cases}$$

Thus, $f(\infty) = \infty$ if and only if $c \le 0$. Notice that in one-dimensional case, if $\gamma \equiv 0$, then

$$g(0) < \infty \Leftrightarrow \lim_{r \to 0} \int_r^1 \frac{s - r}{\alpha(s)} ds < \infty. \tag{4.15}$$

In the present case, $\gamma \le 0$ and

$$\lim_{r \to 0} \int_r^1 \frac{s - r}{s} ds = 1 < \infty.$$

By Theorem 4.2, we conclude that the coupling is successful for all $c \le 0$.

Since $\inf_{r > 0} \alpha(r) = 0$, Corollary 4.3 is not available for this example.

EXAMPLE 4.16 (One-dimensional linear growth model).

$$\sigma(x) = ax + b, \qquad b(x) = cx + d, \qquad a \ne 0.$$

Consider the basic coupling

$$c(x, y) = (ax + b)(ay + b).$$

Then $\gamma(u) = \gamma_*(u) = -2c/a^2$, $\alpha(u) = \alpha^*(u) = a^2u^2$.

It is easy to check either $f(\infty) < \infty$ or $g(0) = \infty$. Hence the coupling is always not successful. Even if $c \le 0$, then $f(\infty) = \infty$ and $f(0) > -\infty$. Hence, it is not difficult to prove that

$$P^{x, y}\Big(\lim_{t \to \infty} |X_t - Y_t| = 0\Big) = 1,$$

but we still have

$$P^{x, y}(T = \infty) = 1, \qquad x \ne y.$$

The above example shows that the basic coupling is useless for the V-metric. However, for negative c, the basic coupling is not only effective but also provides an exponential rate for the W_1-metric (Theorem 2.3). Conversely, for B.M., the basic coupling gives us

$$P^{x, y}(X_t - Y_t = x - y) = 1$$

and so is useless for the W_1-metric. But as we have seen in Example 4.11, we still have an effective coupling for the V-metric. Thus, the suitable couplings are different for different metrics. For different models, we even need different metrics.

Now, we return to the third coupling given in Example 1.2.

EXAMPLE 4.17 (B.M. in $\mathbb{R}^d$).

$$c_{ij}(x, y) = \left(1 - \frac{\alpha|x_i - y_i|}{\beta + |x_i - y_i|}\right)\delta_{ij}, \qquad 1 \le i, j \le d.$$

Observe

$$\begin{aligned}\bar{A}(x, y) &= 2\alpha \sum_{i=1}^{d} \frac{|x_i - y_i|^3}{\beta + |x_i - y_i|} \Big/ |x - y|^2 \\ &\geq \frac{2\alpha d}{\beta + u}\left(\frac{1}{d}\sum_{i=1}^{d} |x_i - y_i|^3\right) \Big/ |x - y|^2 \\ &\geq \frac{2\alpha}{\sqrt{d}} \frac{u}{\beta + u}, \quad \text{if } |x - y| = u, \\ \operatorname{tr} A(x, y) &= \sum_{i=1}^{d} \frac{2\alpha |x_i - y_i|}{\beta + |x_i - y_i|} \\ &= 2\alpha\left(d - \beta \sum_{i=1}^{d} \frac{1}{\beta + |x_i - y_i|}\right) \\ &\leq 2\alpha d\left(1 - \frac{\beta}{\beta + u}\right) \\ &= \frac{2\alpha d u}{\beta + u}, \quad \text{if } |x - y| = u.\end{aligned}$$

Thus,

$$\frac{\operatorname{tr} A(x, y)}{\bar{A}(x, y)} - 1 \leq d - 1,$$

and so

$$\gamma(u) = d - 1, \qquad \alpha(u) = \frac{2\alpha}{\sqrt{d}} \frac{u}{\beta + u}.$$

Our criterion (Theorem 4.2) is available only for $d = 1$.

However, this coupling is successful in any dimension. The reason is that we can use the following simple result to reduce the general case to the case that $d = 1$.

DECOMPOSITION LEMMA 4.18. *If a coupling consists of two independent parts, and each part has the property that when they hit they will move together, then*

$$T = T_1 \vee T_2,$$

where T_1 and T_2 are, respectively, the coupling times of the two parts. In other words, the coupling is successful if and only if each part is successful.

We have seen that Theorem 4.2 is less and less effective as the dimension increases. The role of Lemma 4.18 is to deduce the higher dimensional case to the lower dimensional case. The idea is that, if the components of the original process are independent, we may construct a coupling in two steps: First, for each component, construct a coupling such that after the marginals of the component meet each other, they move together (cf. Theorem 3.1 and Lemma 3.4). Second, link these individual couplings together independently. Example 4.17 illustrates such a construction. As another application of this idea, let us again consider B.M. in $\mathbb{R}^d$. Take the coupling diffusion coefficient as

$$a(t,x,y) = \begin{pmatrix} I & c(t,x,y) \\ c(t,x,y) & I \end{pmatrix},$$

where

$$c_{ij}(t,x,y) = \left(-I_{[0,T_i)}(t) + I_{[T_i,\infty)}(t)\right)\delta_{ij}, \qquad 1 \le i, j \le d,$$

and

$$T_i = \inf\{t \ge 0\colon X_i(t) = Y_i(t)\}, \qquad 1 \le i \le d.$$

This construction works also for the higher dimensional analogue of Example 4.14.

REMARK 4.19. We now would like to know what coupling is the "optimal" for the V-metric. For simplicity, we ignore the drifts for the moment. Based on Theorem 4.2, we may say that a coupling is V-optimal if

(4.20) $\quad a(x,y)$ is nonnegative definite and $\overline{A}(x,y) \neq 0$,

(4.21) $\quad \operatorname{tr} A(x,y) - \overline{A}(x,y)$ achieves the minimum and

(4.22) $\quad \overline{A}(x,y)$ achieves the maximum.

By the Schwarz inequality, we have

$$\operatorname{tr} A(x,y) \ge \overline{A}(x,y). \tag{4.23}$$

Thus, a special case of (4.21) is that (4.23) becomes equality. This happens if and only if

$$c(x,y) + c(x,y)^* = a_1(x) + a_2(y) - \lambda(x,y)^2 \frac{(x-y)(x-y)^*}{|x-y|^2}. \tag{4.24}$$

For B.M., any $\lambda(x,y)$ satisfying

$$0 < \lambda(x,y)^2 \le 4$$

will give us a solution to (4.20) and (4.21). Furthermore, if we assume that $c = c^*$, then the coupling by reflection [i.e., $\lambda(x,y)^2 \equiv 4$] is V-optimal. Next, if $a_1 = a_2 = \sigma^2 =$ constant, then the couplings either by reflection or by projection do satisfy (4.20) and (4.21). If we insist on choosing an orthogonal matrix H mentioned in Remark 2.11, then, for constant $a_1 = a_2 = \sigma^2$, $\det \sigma \neq 0$,

$$H = I - 2\sigma^{-1}(x-y)(x-y)^* \big/ \left|\sigma^{-1}(x-y)\right|^2$$

is a solution to (4.20) and (4.21), but this is no longer true when the matrix σ

depends on x. Similarly, if we consider projection matrix H, the solution is

$$H = I - \sigma^{-1}(x-y)(x-y)^* \big/ \left|\sigma^{-1}(x-y)\right|^2.$$

It is still an open problem to give a general formula for the optimal couplings for V-metric.

5. Rates of convergence in total variation norm. Let us begin this section with an example [Lindvall and Rogers (1986)]. Consider the coupling of B.M. in $\mathbb{R}^d$ by reflection. Using the functional

$$Z_t^\alpha = \exp\left[\alpha(|x-y| - |X_t - Y_t|) - 2\alpha^2 t\right], \qquad \alpha > 0,$$

it is easy to prove that

$$E^{x,y}\exp[-\lambda T] = \exp\left[-\sqrt{\lambda/2}\,|x-y|\right], \qquad \lambda > 0$$

[see Williams (1979), pages 85–86, for example]. Hence

$$P^{x,y}[T > t] = \sqrt{\frac{2}{\pi}}\int_0^{|x-y|/(2\sqrt{t})}\exp\left[-\frac{u^2}{2}\right]du.$$

Thus

$$\begin{aligned}\tfrac{1}{2}\|P(t,x,\cdot) - P(t,y,\cdot)\|_{\mathrm{Var}} &\\ = V(P(t,x,\cdot),P(t,y,\cdot)) &\le E^{x,y}\left[I_{[X_t \ne Y_t]}\right]\\ = P^{x,y}[T>t] &\le \text{Const.}|x-y|/\sqrt{t} \to 0, \quad \text{as } t\uparrow\infty.\end{aligned}$$

On the other hand, it is known that

$$\frac{1}{2}\|P(t,x,\cdot) - P(t,y,\cdot)\|_{\mathrm{Var}} = \sqrt{\frac{2}{\pi}}\int_0^{|x-y|/(2\sqrt{t})}\exp\left[-\frac{u^2}{2}\right]du;$$

thus, the coupling by reflection is exact for the V-metric. Similarly, one can easily check that the coupling by projection will give us the same rate $1/\sqrt{t}$ up to a constant. This procedure produces some estimates for the rates of convergence in some special cases. Now, we are going to use a different idea.

Obviously, if $E(T^m) < \infty$, then

$$t^m V(P(t,x,\cdot),P(t,y,\cdot)) \le t^m P[T>t] \le E[T^m\colon T>t] \to 0, \qquad t\to\infty.$$

This leads us to study the moments of T.

Recall that

$$F_{n,N}(r) = -\int_{1/n}^r C(s)^{-1}\,ds\int_s^N \frac{C(u)}{\alpha(u)}\,du, \qquad 1/n \le r \le N,$$

and define

$$F(r) = \lim_{n\to\infty}\lim_{N\to\infty} F_{n,N}(r), \qquad 0 < r < \infty,$$

$$M_n(s_1,s_2) = \left(C_*(s_1)C(s_2)\right)^{-1}\int_{1/n}^{s_2}\frac{C(u)}{\alpha^*(u)}\,du, \qquad s_1, s_2 > 0.$$

For the following result, we are again comparing with a radial process.

THEOREM 5.1. *Put $r = |x - y|$.*

(i) *If $F(r) > -\infty$, then $E^{x,y}(T) < \infty$.*

(ii) *If $T_{n,N} < \infty$, $P^{x,y}$-a.s. and*

$$\lim_{\substack{n\to\infty \\ N\to\infty}} \iint_{\substack{1/n \le s_1 \le r \\ r \le s_2 \le N}} [M_n(s_1, s_2) - M_n(s_2, s_1)]\, ds_1\, ds_2 \Big/ \int_{1/n}^{N} C_*(s)^{-1}\, ds = +\infty,$$

then $E^{x,y}(T) = \infty$, $x \neq y$.

PROOF. (i) From (4.7), we see that

$$E^{x,y}(T_{n,N}) \le -2F_{n,N}(r). \tag{5.2}$$

Let $N \to \infty$ and then $n \to \infty$ to get

$$E^{x,y}(T) \le -2F(r) < \infty.$$

(ii) Set

$$G_n(r) = \int_{1/n}^{r} C(s)^{-1}\, ds \int_{1/n}^{s} \frac{C(u)}{\alpha^*(u)}\, du.$$

Since $G_n(\rho) \ge 0$, $G_n'(\rho) \ge 0$, $G_n''(\rho) + (1/\rho)\gamma(\rho)G_n'(\rho) = 1/\alpha^*(\rho)$, for $\rho \ge 1/n$. We have

$$2LG_n(\rho(x, y)) \le 1.$$

Hence

$$E^{x,y}G_n\big(\rho\big(X_{T_{n,N}}, Y_{T_{n,N}}\big)\big) \le G_n(r) + \tfrac{1}{2}E^{x,y}(T_{n,N}). \tag{5.3}$$

On the other hand, if we set

$$H_n(r) = \int_{1/n}^{r} C_*(s)^{-1}\, ds,$$

then

$$H_n(\rho) \ge 0, \qquad H_n'(\rho) \ge 0, \qquad H_n''(\rho) + \frac{1}{\rho}\gamma_*(\rho)H_n'(\rho) = 0, \qquad \rho \ge \frac{1}{n}.$$

Because

$$E^{x,y}H_n\big(\rho\big(X_{T_{n,N}}, Y_{T_{n,N}}\big)\big) = H_n(r) + E^{x,y}\int_0^{T_{n,N}} LH_n(\rho(X_u, Y_u))\, du \ge H_n(r),$$

we get

$$P^{x,y}(S_N < T_n) \ge \frac{H_n(r)}{H_n(N)}. \tag{5.4}$$

Combining (5.3) with (5.4), we obtain

$$E^{x,y}(T_{n,N}) \ge 2\frac{H_n(r)G_n(N) - H_n(N)G_n(r)}{H_n(N)}.$$

Since $T_{n,N}$ is increasing as $n \to \infty$ or $N \to \infty$, $T_{n,N} \uparrow T$. Thus, the assumption of the theorem implies that

$$E^{x,y}(T) = \infty. \qquad \square$$

EXAMPLE 5.5 (O.U. process).

$$\sigma(x) = I, \qquad b(x) = -x.$$

Using the coupling by reflection, we get

$$\alpha(r) = \alpha^*(r) = 4,$$

$$\gamma(r) = \gamma_*(r) = -\frac{1}{2}r^2,$$

$$C(r) = \exp\left\{\int_1^r \frac{\gamma(u)}{u}\,du\right\} = \exp\left\{-\frac{1}{4}(r^2-1)\right\},$$

$$-F_{n,N}(r) = \int_{1/n}^r \exp\left[\frac{1}{4}(s^2-1)\right] ds \int_s^N \frac{1}{4}\exp\left\{-\frac{1}{4}(u^2-1)\right\} du$$

$$\le \frac{1}{4}\left(\int_0^r \exp\left(\frac{1}{4}s^2\right) ds\right)\left(\int_0^\infty \exp\left(-\frac{1}{4}u^2\right) du\right).$$

Hence $-F(r) < \infty$ for all $r \in (0,\infty)$, and so

$$E^{x,y}(T) < \infty.$$

EXAMPLE 5.6 (The coupling of B.M. in $\mathbb{R}^d$ by reflection).

$$E^{x,y}(T) = \infty, \qquad x \ne y.$$

Now we investigate the higher moments of the coupling time T.

THEOREM 5.7. *If there exist constants* $\beta > 0$, $0 < \alpha < \beta$, *and* $c = c(\alpha)$ *such that*

$$(5.8) \qquad (\beta - 2)\bar{A}(x,y) + \operatorname{tr} A(x,y) + 2\hat{B}(x,y) \le 0$$

for all (x,y): $0 < \rho(x,y) < \infty$, *and*

$$(5.9) \qquad |F(r)| \le cr^\alpha, \qquad 0 < r < \infty,$$

then

$$(5.10) \qquad E^{x,y}(T^m) < \infty, \qquad m \in [0, \beta/\alpha).$$

PROOF. By (5.8), it is easy to prove that

$$\text{(i)} \qquad \sup_{t\ge 0} E^{x,y}\left(|X(t\wedge T_{n,N}) - Y(t\wedge T_{n,N})|^\beta\right) \le |x-y|^\beta, \qquad n, N \ge 1.$$

Next, by using an integration by parts formula for martingale theory [Stroock and Varadhan (1979), Theorem 1.2.8] and a truncation argument, we may prove that

$$\text{(ii)} \quad E^{x,y}\left(T_{n,N}^{1+m}\right) \le 2m(1+m)E^{x,y}\int_0^{T_{n,N}} |F(|X(s) - Y(s)|)|\, s^{m-1}\,ds.$$

This is the main trick of the proof. Now, by (5.9), Hölder's inequality and (i), we would have

$$E^{x,y}\left[|F(|X(s) - Y(s)|)|; s \le T_{n,N}\right] \le C|x-y|^{\alpha} P^{x,y}\left[T_{n,N} \ge s\right]^{(\beta-\alpha)/\beta}.$$

Inserting this into (ii) and letting $N, n \to \infty$, we would obtain

$$\text{(iii)} \qquad E^{x,y}(T^{1+m}) \le 2Cm(1+m)|x-y|^{\alpha}\int_0^{\infty} s^{m-1}P^{x,y}[T \ge s]^{(\beta-\alpha)/\beta}\,ds.$$

On the other hand, from Theorem 5.1 and 5.9, we see that $E^{x,y}(T) < \infty$. Thus, by using the inequality (iii), we may maximize the number m with property $E^{x,y}(T^m) < \infty$. For more details, refer to the proof of Lemma 7 in Davies (1986). □

EXAMPLE 5.11. Everything is the same as Example 4.12 but for simplicity, we take $a_1 = a_2 = 1$. We know that

$$\gamma(r) = \frac{1}{2}\left(-b_1 r^2 + \frac{b_2}{2}r\right).$$

Hence,

$$C(r) = \exp\left[\int_1^r \frac{\gamma(u)}{u}\,du\right] = \exp\left[-\frac{b_1}{4}(r^2-1) + \frac{b_2}{2}(r-1)\right],$$

$$-F_{n,N}(r) = \frac{1}{4}\int_{1/n}^r \exp\left(\frac{b_1}{4}s^2 - s\right)ds\int_s^N \exp\left(-\frac{b_1}{4}u^2 + \frac{b_2}{2}u\right)du,$$

$$|F(r)| = \frac{1}{4}\int_0^r \exp\left[\frac{b_1}{4}\left(s - \frac{b_2}{b_1}\right)^2\right]ds\int_s^{\infty}\exp\left[-\frac{b_1}{4}\left(u - \frac{b_2}{b_1}\right)^2\right]du$$

and so, for any $0 < \alpha < 1$,

$$\frac{|F(r)|}{r^{\alpha}} \sim \frac{\int_r^{\infty}\exp\left[-(b_1/4)(u - b_2/b_1)^2\right]du}{\alpha r^{\alpha-1}\exp\left[-(b_1/4)(r - b_2/b_1)^2\right]} \to 0, \qquad r \to \infty.$$

Thus (5.10) holds for any $0 < \alpha < 1$, and so

$$E(T^m) < \infty \quad \text{for any } m \ge 0.$$

Finally, we consider exponential estimates for the rate of convergence.

THEOREM 5.12. *Suppose that*

(i) *there exist constants $C \ge 0$, $c > 0$ such that*

$$(5.13) \qquad L\rho^2(x,y) \le C - c\rho^2(x,y),$$

(ii) *there exist $N > N_1 > C/c$ such that*

$$(5.14) \qquad |F_{0,N}(N_1)| = \int_0^N C(s)^{-1}\,ds\int_s^N \frac{C(u)}{\alpha(u)}\,du < \infty$$

and

$$\frac{N_1^2\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} > \frac{C}{c}. \tag{5.15}$$

Then there exists $t_0 > 0$ *such that for* $t \ge t_0$, *we have*

$$\begin{aligned} P^{x,y}(T \ge nt) &\le K_1 k^n, \\ E^{x,y}\rho^2(X_{nt}, Y_{nt}) &\le K_2 k^n, \end{aligned} \tag{5.16}$$

for some constants K_1, $K_2 > 0$ *and* $k \in (0,1)$.

PROOF. Recall

$$f(r) = \int_1^r C(s)^{-1}\,ds, \qquad Lf(\rho(x,y)) \le 0.$$

From (4.9), we see that

$$P^{x,y}(T_n < S_N \wedge t) \ge \frac{f(N) - f(r)}{f(N) - f\left(\frac{1}{n}\right)} - P^{x,y}(T_{n,N} \ge t).$$

Letting $n \to \infty$, we have

$$P^{x,y}(T \le S_N \wedge t) \ge \frac{f(N) - f(r)}{f(N) - f(0)} - P^{x,y}(T_{0,N} \ge t),$$

where $T_{0,N} = T \wedge S_N$, and so

$$P^{x,y}(T \le t) \ge \frac{f(N) - f(r)}{f(N) - f(0)} - P^{x,y}(T_{0,N} \ge t). \tag{5.17}$$

By the condition (5.14) and using (5.2) we get

$$E^{x,y}(T_{0,N}) \le -2F_{0,N}(r).$$

Hence

$$P^{x,y}(T_{0,N} \ge t) \le \frac{E^{x,y}(T_{0,N})}{t} \le \frac{-2F_{0,N}(r)}{t}. \tag{5.18}$$

Combining (5.17) with (5.18) we obtain

$$\begin{aligned} P^{x,y}(T \le t) &\ge \frac{f(N) - f(r)}{f(N) - f(0)} + \frac{2F_{0,N}(r)}{t} \\ &\ge \frac{\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} + \frac{2F_{0,N}(r)}{t} \end{aligned} \tag{5.19}$$

for all (x, y): $0 < \rho(x,y) = r < N_1$.

Let

$$\alpha = \frac{N_1^2\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} - \frac{C}{c}.$$

Then $\alpha > 0$ by (5.15). Clearly, we can find $t_1 > 0$ such that

$$2N_1^2 \frac{|F_{0,N}(N_1)|}{t} \le \frac{\alpha}{2},$$

$$\frac{\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} - \frac{2|F_{0,N}(N_1)|}{t} > 0,$$

for all $t \ge t_1$. Also, we can find $t_2 > 0$ such that

$$\frac{C/c}{1 - e^{-ct}} < \frac{C}{c} + \frac{\alpha}{2}$$

for all $t \ge t_2$. Take $t_0 = t_1 \vee t_2$. Then for all $t \ge t_0$, we have

$$N_1^2 \left[\frac{\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} + \frac{2F_{0,N}(N_1)}{t} \right] > \frac{C/c}{1 - e^{-ct}}. \tag{5.20}$$

Now we fix $t \ge t_0$ and let

$$\frac{\int_{N_1}^N C(s)^{-1}\,ds}{\int_0^N C(s)^{-1}\,ds} - \frac{2F_{0,N}(N_1)}{t} = 1 - \delta, \qquad 0 < \delta < 1.$$

By (5.19), (5.20), (5.13) and Theorem 2.3 we arrive at

(a) $$P^{x,y}(T > t) < \delta, \qquad 0 < |x - y| < N_1,$$

(b) $$N_1^2(1 - \delta)(1 - e^{-ct}) > C/c,$$

(c) $$E^{x,y}\rho^2(X_t, Y_t) \le C/c + e^{-ct}\rho^2(x, y).$$

Let $\tau_n = nt \wedge T$ and $\{P_\omega\}$ be a regular conditional probability distribution $P^{x,y}|\mathscr{M}_{\tau_{n-1}}$. Then $\delta_{(X(\tau_{n-1}), Y(\tau_{n-1}))} \otimes_{\tau_{n-1}} P_\cdot$ is the solution to the martingale problem for the coupling operator $L(a, b)$ starting from $(X(\tau_{n-1}), Y(\tau_{n-1}))$. Define

$$I_n = I_{[\tau_n < T]} = I_{[\tau_n = nt]},$$

$$J_n = \rho^2(X(\tau_n), Y(\tau_n))I_n.$$

Then $I_n \le I_{n-1}$ and $I_{n-1} = 0$ implies $I_n = 0$. Thus, by using (a), we obtain

$$\begin{aligned} E^{x,y}(I_n) &= E^{x,y}\left[I_{n-1}E^{x,y}(I_n|\mathscr{M}_{\tau_{n-1}})\right] \\ &= E^{x,y}\left[I_{n-1}E^{\delta\otimes P_\cdot}(I_n)\right] \\ &= E^{x,y}\left[I_{n-1}E^{\delta\otimes P_\cdot}(I_n), \rho(X(\tau_{n-1}), Y(\tau_{n-1})) \le N_1\right] \\ &\quad + E^{x,y}\left[I_{n-1}E^{\delta\otimes P_\cdot}(I_n), \rho(X(\tau_{n-1}), Y(\tau_{n-1})) > N_1\right] \\ &\le \delta E^{x,y}(I_{n-1}) + \frac{1}{N_1^2}E^{x,y}(J_{n-1}), \end{aligned}$$

where $\delta \otimes P_\cdot = \delta_{(X(\tau_{n-1}), Y(\tau_{n-1}))} \otimes_{\tau_{n-1}} P_\cdot$. Next, by using (c), we get

$$E^{x,y}(J_n) = E^{x,y}\left[I_{n-1}E^{x,y}(J_n|\mathscr{M}_{\tau_{n-1}})\right] \le (C/c)E^{x,y}(I_{n-1}) + e^{-ct}E_{x,y}(J_{n-1}).$$

Finally, the assertion (b) guarantees that the eigenvalues of the matrix

$$\begin{pmatrix} \delta & 1/N_1^2 \\ C/c & e^{-ct} \end{pmatrix}$$

are less than 1. Let λ_1, λ_2 be the eigenvalues and take $k \in [\lambda_1 \vee \lambda_2, 1)$. Then (5.16) holds for some $K_1, K_2 > 0$. □

EXAMPLE 5.21 (O.U. process).

$$\begin{aligned} \sigma(x) &= I, \qquad b(x) = -x, \\ L\rho^2(x, y) &= 4 - 2\rho^2(x, y), \qquad C = 4, \qquad c = 2, \\ C(r) &= \exp\{-(r^2 - 1)/4\}, \\ -F_{0,N}(r) &= \int_0^r e^{s^2/4}\, ds \int_s^N \tfrac{1}{4} e^{-u^2/4}\, du < \infty, \\ f(r) &= \int_1^r \exp[(s^2 - 1)/4]\, ds. \end{aligned}$$

Take $N_1 > 2 = C/c$. Then $2/N_1^2 < 1$. Since

$$\lim_{N \to \infty} f(N) = \infty$$

for fixed N_1 and

$$\lim_{N \to \infty} \frac{f(N) - f(N_1)}{f(N) - f(0)} = 1,$$

we can choose N large enough such that

$$\frac{f(N) - f(N_1)}{f(N) - f(0)} > \frac{2}{N_1^2}.$$

This implies (5.15) and hence the hypotheses of Theorem 5.12 are satisfied.

Acknowledgments. This work was done when the authors were visiting the Department of Mathematics, University of Edinburgh. The authors would like to thank the Department for its hospitality, especially Professor T. J. Lyons. The first author also acknowledges the support of the S.E.R.C. of the United Kingdom. Finally, thanks are given to a referee for his detailed comments on the earlier version of this paper.

REFERENCES

BASIS, V. YA. (1980). Stationarity and ergodicity of Markov interacting processes. In *Multicomponent Random Systems* (R. L. Dobrushin and Ya. G. Sinai, eds.) 37–58. Dekker, New York.

BHATTACHARYA, R. N. and RAMASUBRAMANIAN, S. (1982). Recurrence and ergodicity of diffusions. *J. Multivariate Anal.* **12** 95–122.

CHEN, M. F. (1986a). Coupling of jump processes. *Acta Math. Sinica* **2** 123–136.

CHEN, M. F. (1986b). *Jump Processes and Interacting Particle Systems*. Beijing Normal Univ. Press, Beijing. (In Chinese.)

CHEN, M. F. (1987a). Coupling of jump processes. II. *Chin. Ann. Math.* To appear.

CHEN, M. F. (1987b). Existence theorems for interacting particle systems with non-compact state spaces. *Sci. Sinica* (*Ser. A*) **30** 148–156.

DAVIES, P. L. (1986). Rates of convergence to the stationary distribution for k-dimensional diffusion processes. *J. Appl. Probab.* **23** 370–384.

DOBRUSHIN, R. L. (1970). Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.* **15** 458–486.

FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications* **1**. Academic, New York.

GIVENS, G. R. and SHORTT, R. M. (1984). A class of Wasserstein metrics for probability distributions. *Michigan Math. J.* **31** 231–240.

GRIFFEATH, D. (1978). Coupling methods for Markov processes. *Adv. in Math.* **2** 1–43.

IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

KENDALL, W. S. (1986a). Stochastic differential geometry, a coupling property and harmonic maps. *J. London Math. Soc. (2)* **33** 554–566.

KENDALL, W. S. (1986b). Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics* **19** 111–129.

LIGGETT, T. M. (1985). *Interacting Particle Systems*. Springer, New York.

LINDVALL, T. (1983). On coupling of diffusion processes. *J. Appl. Probab.* **20** 82–93.

LINDVALL, T. and ROGERS, L. C. G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872.

STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.

WILLIAMS, D. (1979). *Diffusions, Markov Processes, and Martingales* **1** *Foundations*. Wiley, New York.

DEPARTMENT OF MATHEMATICS
BEIJING NORMAL UNIVERSITY
BEIJING, 100875
PEOPLE'S REPUBLIC OF CHINA

DEPARTMENT OF MATHEMATICS
HENAN TEACHER'S UNIVERSITY
XINXIANG PROVINCE
PEOPLE'S REPUBLIC OF CHINA