

SHY COUPLINGS

Itai Benjamini, Krzysztof Burdzy and Zhen-Qing Chen

Abstract. A pair (X, Y) of Markov processes is called a Markov coupling if X and Y have the same transition probabilities and (X, Y) is a Markov process. We say that a coupling is “shy” if there exists a (random) $\varepsilon > 0$ such that $\text{dist}(X_t, Y_t) > \varepsilon$ for all $t \geq 0$. We investigate whether shy couplings exist for several classes of Markov processes.

1. Introduction. The proofs of the main theorems in two recent papers, [BC2] and [BCJ], contained arguments showing that certain processes come arbitrarily close to each other, at least from time to time, as time goes to infinity, with probability one. The proofs were based on ideas specific to the models and were rather tedious. We decided to examine several classes of Markov processes in order to determine the conditions under which there exists a pair of Markov processes defined on the same probability space such that each marginal process has the same transition probabilities and the two processes do not come close to each other at any time. Although we do not have a complete solution to this problem, we offer a number of results whose diversity points to a rich theory. Some of our theorems, examples and techniques may have interest of their own.

We will focus on two classes of processes—reflected Brownian motions on Euclidean domains and Brownian motions on graphs. The second class of processes is really discrete in nature, in the sense that similar techniques work for random walks on graphs. We chose these classes of processes because similar processes appeared in our research in the past.

For a general overview of coupling techniques, see [L].

The rest of the paper is organized as follows. We present basic definitions and elementary examples in Section 2. Section 3 is devoted to Brownian motions on graphs. We show that there exists a shy coupling for Brownian motions on a graph if all its vertices have degree 3 or higher. Four examples are also given to illustrate the case when the graph has some vertices of degree one. Section 4 deals with reflected Brownian motions on Euclidean

Research partially supported by NSF grant DMS-0303310 (KB and ZC).

domains, showing that there exist no shy couplings on C^1 -smooth bounded strictly convex domains.

2. Preliminaries and elementary examples. Unless specified otherwise, all pairs of processes (X, Y) considered in this paper will be “Markov couplings,” i.e., they will satisfy the following assumptions.

- (i) $\{X_t, t \geq 0\}$, $\{Y_t, t \geq 0\}$ and $\{(X_t, Y_t), t \geq 0\}$ are Markov, and the transition probabilities for X and Y are identical.
- (ii) The distribution of $\{X_t, t \geq s\}$ conditional on $\{(X_s, Y_s) = (x, y)\}$ is the same as the distribution of $\{X_t, t \geq s\}$ conditional on $\{X_s = x\}$, for all x, y and s .

Our definition of a Markov coupling is slightly different from similar concepts in the literature. One could investigate the question of whether our results hold for “couplings” defined in other ways, for example, whether condition (ii) is essential. However, we feel that there are more exciting open problems in this area—see the end of Section 4.

The following elementary discrete-time example shows that there exist couplings that satisfy (i) but do not satisfy (ii).

Example 2.1. We take $\{0, 1\}$ as the state space of a discrete time Markov process and we let $\{X_k, k \geq 0\}$ be a sequence of i.i.d. random variables with $\mathbf{P}(X_k = 0) = \mathbf{P}(X_k = 1) = 1/2$. We will define a process $\{Y_k, k \geq 0\}$ with the same distribution as $\{X_k, k \geq 0\}$. We let Y_0 be independent of $\{X_k, k \geq 0\}$. For $k \geq 1$, we construct Y_k so that $\mathbf{P}(Y_k = 0 \mid X_{k-1} = 0) = 0.7$ and $\mathbf{P}(Y_k = 1 \mid X_{k-1} = 1) = 0.7$. Moreover, for every $k \geq 1$, we make Y_k independent of X_j 's for $j < k - 1$. It is elementary to check that $\{X_k, k \geq 0\}$, $\{Y_k, k \geq 0\}$ and $\{(X_k, Y_k), k \geq 0\}$ are Markov but for $j \geq 1$, the distribution of $\{Y_k, k \geq j\}$ conditional on $\{Y_{j-1} = 0\}$ is not the same as the distribution of $\{Y_k, k \geq j\}$ conditional on $\{(X_{j-1}, Y_{j-1}) = (0, 0)\}$.

We will assume that the state space \mathcal{S} for Markov processes X and Y is metric and we will let \mathbf{d} denote the metric. The open ball with center x and radius r will be denoted $\mathcal{B}(x, r)$. The shortest path between two points in \mathcal{S} will be called a geodesic. For some pairs of points, there may be more than one geodesic joining them.

Definition 2.2. A coupling (X, Y) will be called shy if one can find two distinct points x and y in the state space with

$$\mathbf{P} \left(\inf_{0 \leq t < \infty} \mathbf{d}(X_t, Y_t) > 0 \mid X_0 = x, Y_0 = y \right) > 0.$$

Note that the term “shy coupling” is a label for a family of Markov transition probabilities.

We proceed with completely elementary examples of shy and non-shy couplings.

Examples 2.3. (i) Let X be a Brownian motion in \mathbb{R}^d and let $0 \neq y \in \mathbb{R}^d$ be a fixed vector. Let $Y_t = X_t + y$ for all $t \geq 0$. Then (X, Y) is a shy coupling.

(ii) Let X be a Brownian motion on the unit circle in \mathbb{R}^2 and let $\theta \in (0, 2\pi)$ be a fixed number. We define Y_t using complex notation, $Y_t = e^{i\theta} X_t$ for all $t \geq 0$. Then (X, Y) is a shy coupling.

(iii) The last two examples can be easily generalized to a wide class of Markov processes on spaces \mathcal{S} with a group structure. If there is a group element $a \neq 0$ such that $a + X$ is a Markov process having the same transition probabilities as X and $\inf_{b \in \mathcal{S}} \mathbf{d}(b, a + b) > 0$ then $(X, a + X)$ is a shy coupling.

(iv) Let X and Y be independent Brownian motions in \mathbb{R}^d . Then (X, Y) is a shy coupling if and only if $d \geq 3$.

In the sequel, for $a, b \in \mathbb{R}$, $a \wedge b := \min\{a, b\}$, $a \vee b := \max\{a, b\}$ and $a^+ := a \vee 0$. For $a > 0$, $[a]$ denotes the largest integer that does not exceed a .

3. Brownian motion on graphs. In this section, we will consider processes whose state space is a finite or infinite graph. More precisely, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph, where \mathcal{V} is the set of vertices and \mathcal{E} is the set of edges. We will assume that all vertices have a finite degree, i.e., for every vertex there are only a finite number of edges emanating from this vertex, but we do not assume that this number is bounded over the set of all vertices. We allow an edge to have both endpoints attached to one vertex. Every vertex will be attached to at least one edge. We will identify edges with finite open line segments (connected subsets of \mathbb{R}), with finite and strictly positive length, and we will identify vertices with topological endpoints of edges. In this way, we can identify the graph \mathcal{G} with a metric space $(\mathcal{S}, \mathbf{d})$, where $\mathcal{S} = \mathcal{E} \cup \mathcal{V}$, and $\mathbf{d}(x, y)$ is the shortest path between x and y along the edges of the graph. We will assume that the length of any edge is bounded below by $r_0 > 0$.

Next we will construct “Brownian motion” X on \mathcal{S} . See [FW] for a definition of a general diffusion on a graph. We leave it to the reader to check that our somewhat informal description of the process is consistent with the rigorous construction given in [FW]. By assumption, our process will be strong Markov. Suppose that $x \in e \in \mathcal{E}$ and x is not an endpoint of e . Recall that e can be identified with a line segment, say, $e = [0, y]$. Then $x \in (0, y)$. If $X_0 = x$, then the process X evolves just like the standard one-dimensional

Brownian motion until the exit time from $(0, y)$. Next suppose that $x \in \mathcal{S}$ is a vertex. Then there are n edges e_1, e_2, \dots, e_n , attached to x , with $n \geq 1$. Choose a small $r > 0$ such that the ball $\mathcal{B}(x, r)$ consists of line segments I_j , $j = 1, 2, \dots, k$, which are disjoint except that they have one common endpoint x . Note that $n \leq k \leq 2n$, but not necessarily $k = n$, because some edges may have both endpoints at x . We will describe the evolution of X starting from x until its exit time from $\mathcal{B}(x, r)$. Generate a reflected Brownian motion R on $[0, \infty)$, starting from 0, and kill it at the first exit time from $[0, r]$, denoted T_r . Label its excursions from 0 with numbers $1, 2, \dots, k$, in such a way that every excursion has a label chosen uniformly from $\{1, \dots, k\}$ and independently of all other labels. Then we define X_t for $t \in [0, T_r]$ so that $\mathbf{d}(X_t, x) = R_t$ and $X_t \in I_j$, where j is the label of the excursion of R from 0 that straddles t (if $R_t = 0$ then obviously $X_t = x$). This defines the process X_t until its exit time from $\mathcal{B}(x, r)$. What we said so far and the strong Markov property uniquely define the distribution of X . Note that when the degree of a vertex x is 1 then X is best described as a process reflected at x . The process X spends zero amount of time at any vertex.

Recall that we have assumed that the length of all edges is bounded below by $r_0 > 0$. Under this assumption, the process cannot visit an infinite number of vertices in a finite amount of time. Hence, the above construction defines a process for all $t \geq 0$. Another consequence of the assumptions that all edges have length greater than r_0 and all vertices have finite degree is that for any two points in \mathcal{S} there is only a finite number of geodesics joining them. It is clear from our construction that vertices of degree 2 will play no essential role in the paper and can be ignored. So we will assume without loss of generality that there are no vertices of degree 2.

Theorem 3.1. *If all vertices of \mathcal{G} have degree 3 or higher then there exists a shy coupling for Brownian motions on \mathcal{S} .*

Proof. We will construct a coupling (X, Y) of Brownian motions on \mathcal{S} such that X and Y move in an independent way when they are far apart and they move in a “synchronous” way when they are close together. Clearly, independent processes do not form a shy coupling on a finite graph. Remark 3.2 below explains why it is hard, perhaps impossible, to construct a “synchronous” shy coupling.

For any $x, y \in \mathcal{S}$ with $\mathbf{d}(x, y) > r_0/4$, we will define (X_t, Y_t) starting from $(X_0, Y_0) = (x, y)$, for $t \in [0, \tau]$, where τ is a random time depending on x and y . Then we will explain how one can define (X_t, Y_t) for $t \in [0, \infty)$ by pasting together different pieces of

the trajectory.

(i) Recall that the length of any edge is at least $r_0 > 0$. First suppose that $\mathbf{d}(x, y) \geq 3r_0/4$. Then we let $\{(X_t, Y_t), t \in [0, \tau]\}$ be two independent copies of Brownian motion on \mathcal{S} and we let $\tau = \inf\{t > 0 : \mathbf{d}(X_t, Y_t) = r_0/2\}$.

(ii) Next suppose that $x, y \in \mathcal{S}$ are such that $\mathbf{d}(x, y) \in (r_0/4, 3r_0/4)$, and none of these points is a vertex. Let

$$\sigma(r) = \frac{(4|r| - r_0)^+}{r_0} \wedge 1, \quad (3.1)$$

and B and B' be independent Brownian motions on \mathbb{R} starting from the origin. Let $U_t = B_t$ and

$$dV_t = \sqrt{1 - \sigma^2(U_t - V_t)}dB_t + \sigma(U_t - V_t)dB'_t, \quad (3.2)$$

with $V_0 = v_0 = \mathbf{d}(x, y) > r_0/4$. Then, if we write $Z_t = V_t - U_t$, we obtain

$$Z_t = v_0 + \int_0^t (\sqrt{1 - \sigma^2(Z_s)} - 1)dB_s + \int_0^t \sigma(Z_s)dB'_s,$$

and for $Z'_t \stackrel{\text{df}}{=} Z_t - r_0/4$,

$$Z'_t = v_0 - r_0/4 + \int_0^t (\sqrt{1 - \sigma^2(Z'_s + r_0/4)} - 1)dB_s + \int_0^t \sigma(Z'_s + r_0/4)dB'_s.$$

So

$$Z'_t = v_0 - r_0/4 + \int_0^t \gamma(Z'_s)dW_s,$$

where

$$\gamma(r) \stackrel{\text{df}}{=} (\sqrt{1 - \sigma^2(r + r_0/4)} - 1)^2 + \sigma^2(r + r_0/4)$$

and W is a Brownian motion on \mathbb{R} with $W_0 = 0$. The process Z' has the same distribution as

$$t \mapsto v_0 - r_0/4 + W_{\tau_t},$$

where

$$\tau_t := \inf \left\{ s > 0 : \int_0^s \gamma(v_0 - r_0/4 + W_s)^{-2} ds > t \right\}.$$

Note that for small $r > 0$, $\gamma(r) = O(r^2)$ and so in particular $\int_{0+} \gamma(r)^{-2} dr = \infty$. Thus by Lemma V.5.2 of [KS],

$$\int_0^{\tau_t} \gamma(v_0 - r_0/4 + W_s)^{-2} ds = \infty$$

almost surely, where $T_0 = \inf\{t > 0 : v_0 - r_0/4 + W_s = 0\}$. We conclude that Z' never hits 0; in other words, Z never reaches $r_0/4$.

Suppose that $X_0 = x$, $Y_0 = y$ and recall that we have assumed that $\mathbf{d}(x, y) \in (r_0/4, 3r_0/4)$. Suppose that $x \in e_1 \in \mathcal{E}$ and $y \in e_2 \in \mathcal{E}$. Let $e_1 \setminus \{x\}$ consist of two line segments e_1^ℓ and e_1^r , with e_1^r being the one closer to y . Similarly, $e_2 \setminus \{y\}$ consists of two line segments e_2^ℓ and e_2^r , and e_2^ℓ is closer to x . We will define X and Y on an interval $[0, \tau]$ to be specified later. We define X_t on e_1 to be such that $\mathbf{d}(X_t, x) = |U_t|$ and $X_t \in e_1^\ell$ if and only if $U_t < 0$. We define the process Y_t on e_2 by conditions $\mathbf{d}(Y_t, y) = |V_t - v_0|$ and $Y_t \in e_2^\ell$ if and only if $V_t < v_0$. We let τ be the first time $t > 0$ that X_t or Y_t is at a vertex, or $\mathbf{d}(X_t, Y_t) = 3r_0/4$. We see that over the interval $[0, \tau)$, the distance between X_t and Y_t remains in the interval $(r_0/4, 3r_0/4)$.

(iii) This part of our argument is based on the “skew Brownian motion.” The skew Brownian motion U is a real-valued diffusion which satisfies the stochastic differential equation

$$U_t = B_t + \beta L_t^U, \quad (3.3)$$

where B is a given Brownian motion with $B_0 = 0$, $\beta \in [-1, 1]$ is a fixed constant and L^U is the symmetric local time of U at 0, i.e.,

$$L_t^U = \lim_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{(-\varepsilon, \varepsilon)}(U_s) ds. \quad (3.4)$$

The existence and uniqueness of a strong solution to (3.3)-(3.4) was proved in [HS]. In the special case of $\beta = 1$, the solution to (3.3) is the reflected Brownian motion. An alternative way to define the skew Brownian motion is the following. Consider the case $\beta > 0$. Take a standard Brownian motion B_t' and flip every excursion of B_t' from 0 to the positive side with probability β , independent of what happens to other excursions (if an excursion is on the positive side, it remains unchanged). The resulting process has the same distribution as U defined by (3.3)-(3.4). For more information and references, see recent papers on skew Brownian motion, [BC1] and [BK].

Suppose that $x, y \in \mathcal{S}$, $\mathbf{d}(x, y) \in (r_0/4, 3r_0/4)$, and x is a vertex. Note that y is not a vertex. By assumption, the degree k of vertex x is 3 or greater. Suppose that B_t is a Brownian motion on \mathbb{R} and let U be a solution to (3.3)-(3.4), with β defined by $(1 - \beta)/(1 + \beta) = k - 1$. Note that $\beta < 0$. We label negative excursions of U from 0 with numbers $1, 2, \dots, k - 1$, in such a way that every excursion has a label chosen uniformly from this set and independently of all other labels.

Suppose that B' is a Brownian motion independent of B . Recall the definition of the function σ and the process V given in (3.1) and (3.2), respectively, with $V_0 = v_0 = \mathbf{d}(x, y)$. Then, if we write $Z_t = V_t - U_t$, we obtain

$$Z_t = v_0 - \beta L_t^U + \int_0^t (\sqrt{1 - \sigma^2(Z_s)} - 1) dB_s + \int_0^t \sigma(Z_s) dB'_s,$$

and for $Z'_t \stackrel{\text{df}}{=} Z_t - r_0/4$,

$$Z'_t = v_0 - r_0/4 - \beta L_t^U + \int_0^t (\sqrt{1 - \sigma^2(Z'_s + r_0/4)} - 1) dB_s + \int_0^t \sigma(Z'_s + r_0/4) dB'_s.$$

We have already pointed out that for small $r > 0$,

$$\gamma(r) \stackrel{\text{df}}{=} (\sqrt{1 - \sigma^2(r + r_0/4)} - 1)^2 + \sigma^2(r + r_0/4) = O(r^2).$$

Since $\beta < 0$, the process $-\beta L_t^U$ is nondecreasing. These observations and the argument used in the first half of (ii) imply that Z' never hits 0, i.e., Z never reaches $r_0/4$.

The ball $\mathcal{B}(x, 3r_0/4)$ consists of line segments I_j , $j = 1, 2, \dots, k$. We assume that I_k is the line segment containing y . We define X_t on these line segments so that $\mathbf{d}(X_t, x) = |U_t|$. If $U_t > 0$ then $X_t \in I_k$. If $U_t < 0$ then $X_t \in I_j$, where j is the label of the excursion of U straddling t . Suppose that $y \in e \in \mathcal{E}$. Let $e \setminus \{y\}$ consist of two line segments e^ℓ and e^r , with e^ℓ being the one closer to x . We define Y_t on e by $\mathbf{d}(Y_t, y) = |V_t - v_0|$. We let $Y_t \in e^\ell$ if and only if $V_t < v_0$. We let τ be the infimum of t such that Y_t is at a vertex, or $\mathbf{d}(X_t, Y_t) = 3r_0/4$. Observe that over the interval $[0, \tau)$, the distance between X_t and Y_t remains in the interval $(r_0/4, 3r_0/4)$.

Now we will define the process (X_t, Y_t) for all $t \geq 0$, assuming that $X_0 = x$, $Y_0 = y$ and $\mathbf{d}(x, y) > r_0/4$. We use one of the parts (i)-(iii) of the proof to define the process (X, Y) on an interval $[0, \tau_1]$. Then we proceed by induction. Suppose that the process has been defined on an interval $[0, \tau_k]$ and $\mathbf{d}(X_{\tau_k}, Y_{\tau_k}) > r_0/4$. Then we use the appropriate part (i)-(iii) of the proof to extend the process, using the strong Markov property at τ_k , to an interval $[0, \tau_{k+1}]$. It is easy to see that $\tau_k \rightarrow \infty$ a.s., so the process (X_t, Y_t) is defined for all $t \geq 0$. It is straightforward to check that $\{X_t, t \geq 0\}$ and $\{Y_t, t \geq 0\}$ are Brownian motions on \mathcal{S} and (X, Y) is a shy Markov coupling, as defined in Section 2. \square

Remark 3.2. One may wonder whether it is possible to construct the shy coupling in the proof of Theorem 3.1 using the skew Brownian motion in such a way that the distance

between X and Y does not change on time intervals where both processes stay away from vertices (“synchronous coupling”). This idea works well for many graphs but runs into technical problems when we have a configuration similar to that in Fig. 3.1, with many geodesics joining two vertices. Suppose that the lengths of edges in Fig. 3.1 are chosen so that there are 6 geodesics between x and y . We will argue that if $X_0 = x$ and $Y_0 = y$ then for small $t > 0$ the distance between X and Y has to decrease. This is because X will have to move towards z with probability $5/7$, and Y will have to move towards z with probability $1/3$. Since $5/7 + 1/3 > 1$, X and Y will find themselves on a geodesic from x to y , moving away from their starting points towards each other, with positive probability. For this reason, we could not find a “synchronous” shy coupling based on the skew Brownian motion.

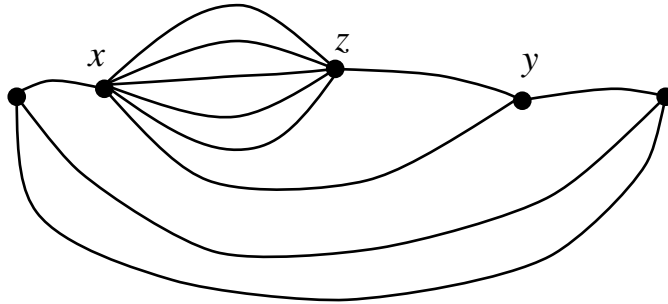


Figure 3.1.

The rest of this section is devoted to graphs that have at least one vertex of degree 1 (recall that vertices of degree 2 can be ignored and we assume that \mathcal{G} does not contain any of them). We do not have a general theorem covering all graphs with some vertices of degree 1 but we have four examples illustrating some special cases.

Example 3.3. This example is similar to Examples 2.3 (i)-(iii). Suppose that there exists an isometry $I : \mathcal{S} \rightarrow \mathcal{S}$ such that $\inf_{x \in \mathcal{S}} \mathbf{d}(x, I(x)) > 0$. It is not hard to show that this holds if the isometry has no fixed points, i.e., if there does not exist $x \in \mathcal{S}$ with $I(x) = x$. If such an isometry exists, then we can first construct the process X and then take $Y_t = I(X_t)$ for all $t \geq 0$. Obviously, thus constructed coupling (X, Y) is shy. Fig. 3.2 shows that a graph with some vertices of degree 1 may have this property.

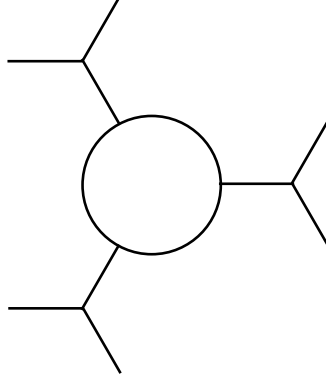


Figure 3.2.

Our next lemma is a large deviations-type estimate. Recall that all edges are at least $r_0 > 0$ units long, by assumption.

Let X be Brownian motion on a graph \mathcal{S} . For $A \subset \mathcal{S}$, define $T_A = \inf\{t \geq 0 : X_t \in A\}$.

Lemma 3.4. *Assume that the degrees of all vertices are bounded above by m_0 . There exist constants $c_0 > 0$, $t_0 < \infty$ depending on r_0 only such that for $t \in (0, t_0)$ and $r > 0$ with $r^2 > t$ and $\mathcal{B}(x, r)^c \neq \emptyset$, we have*

$$\left(\frac{c_0}{m_0} \sqrt{\frac{t}{2\pi}} \frac{1}{2r} \right)^{\lceil r/r_0 \rceil} \exp\left(-\frac{r^2}{2t}\right) \leq \mathbf{P}(T_{\mathcal{B}(x,r)^c} < t) \leq (m_0^{\lceil r/r_0 \rceil})! \sqrt{\frac{2t}{\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right).$$

In applications of the above estimates, r will be “fixed” and then a t (much smaller than r^2) will be chosen. For this reason, we did not try to optimize the non-exponential factors—most likely they are not best possible.

Proof. (i) First we will prove the lower bound. We start with some preliminary estimates.

Suppose that B is a Brownian motion on \mathbb{R} with $B_0 = 0$ and let $T_r^B = \inf\{t \geq 0 : B_t = r\}$. Then $\mathbf{P}(T_r^B < t) = 2 \frac{1}{\sqrt{2\pi t}} \int_r^\infty e^{-u^2/2t} du$ for $r > 0$. The following inequalities are well known (see Problem 9.22 on page 112 of [KS]):

$$\frac{r}{1+r^2} e^{-r^2/2} \leq \int_r^\infty e^{-u^2/2} du \leq \frac{1}{r} e^{-r^2/2}.$$

So for $r \geq 1$,

$$\frac{1}{2r} e^{-r^2/2} \leq \int_r^\infty e^{-u^2/2} du \leq \frac{1}{r} e^{-r^2/2}.$$

By scaling we obtain for $t \leq r^2$,

$$\sqrt{\frac{t}{2\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right) \leq \mathbf{P}(T_r^B < t) = \sqrt{\frac{2}{\pi}} \int_{r/\sqrt{t}}^{\infty} e^{-v^2/2} dv \leq \sqrt{\frac{2t}{\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right). \quad (3.5)$$

For $r \leq 1$, we have a trivial lower bound $\int_1^{\infty} e^{-u^2/2} du \stackrel{\text{df}}{=} c_0 > 0$. For $t \geq r^2$, the same upper bound holds but the lower bound has to be replaced with a trivial bound $c_1 = \sqrt{\frac{2}{\pi}} c_0 > 0$.

Let $t_0 < \infty$ be the largest real such that

$$1 - 2 \exp(-r_0^2/(2t_0)) \geq 1/2.$$

We will derive an estimate for $\mathbf{P}(T_r^B < t \wedge T_{-r_0/2}^B)$, for $r \geq r_0/2$ and $t < t_0$. If $r^2 < t_0$, then $r \leq c_2 r_0$ for some constant $c_2 < \infty$. In this case, it follows easily from the support theorem for Brownian motion that $\mathbf{P}(T_r^B < t \wedge T_{-r_0/2}^B) > c_3 > 0$ for every $t \in (r^2, t_0)$.

Now suppose that $t \leq r^2 \wedge t_0$. If Brownian motion hits $-r_0/2$ and then it reaches r in t seconds or less, it has to go from level $-r_0/2$ to level r in t seconds or less. Hence, by the strong Markov property applied at $T_{-r_0/2}^B$,

$$\begin{aligned} \mathbf{P}(T_r^B < t \wedge T_{-r_0/2}^B) &\geq \mathbf{P}(T_r^B < t) - \mathbf{P}(T_{r+r_0/2}^B < t) \\ &\geq \sqrt{\frac{t}{2\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right) - \sqrt{\frac{2t}{\pi}} \frac{1}{r+r_0/2} \exp\left(-\frac{(r+r_0/2)^2}{2t}\right) \\ &\geq \sqrt{\frac{t}{2\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right) - 2\sqrt{\frac{t}{2\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right) \exp\left(-\frac{r_0^2}{2t}\right) \\ &= \sqrt{\frac{t}{2\pi}} \frac{1}{r} \exp\left(-\frac{r^2}{2t}\right) \left[1 - 2 \exp\left(-\frac{r_0^2}{2t}\right)\right] \\ &\geq \sqrt{\frac{t}{2\pi}} \frac{1}{2r} \exp\left(-\frac{r^2}{2t}\right). \end{aligned} \quad (3.6)$$

Suppose that y_1 is a vertex, $\mathbf{d}(y_1, y_2) = r_0/2$, and y_3 lies between y_1 and y_2 so that $\mathbf{d}(y_1, y_3) + \mathbf{d}(y_3, y_2) = r_0/2$. Let $\mathbf{d}(y_1, y_3) = r_1 \in [0, r_0/2]$ and suppose that $X_0 = y_3$. The process $R_t \stackrel{\text{df}}{=} \mathbf{d}(X_t, y_1)$ is a one-dimensional reflected Brownian motion with $R_0 = r_1$, at least until it reaches r_0 . Let $\{z_1, z_2, \dots, z_k\}$ be the set of all points with $\mathbf{d}(z_j, y_1) = r_0/2$ (y_2 is one of these points). Let $T_{r_0/2}^R = \inf\{t \geq 0 : R_t = r_0/2\}$. If X visited y_1 before $T_{r_0/2}^R$ then it is at y_2 at time $T_{r_0/2}^R$ with probability $1/k$, by symmetry. Since $k \leq m_0$, the probability that X starts at y_3 and reaches y_2 in s seconds or less is greater than or equal to the probability that reflected Brownian motion that starts from r_1 reaches $r_0/2$ in s

seconds or less, divided by m_0 . The last probability is bounded below by the analogous probability for the non-reflected Brownian motion, so using (3.5) we obtain,

$$\mathbf{P}(T_{y_2}^X \leq s \mid X_0 = y_3) \geq \frac{1}{m_0} \sqrt{\frac{s}{2\pi}} \frac{1}{r_0/2 - r_1} \exp\left(-\frac{(r_0/2 - r_1)^2}{2s}\right), \quad (3.7)$$

for $s \leq (r_0/2 - r_1)^2$. If $s \geq (r_0/2 - r_1)^2$, the bound is c_1/m_0 . Note that these estimates hold for all $r_1 \in [0, r_0/2]$, including $r_1 = 0$.

Consider any $x' \in \partial\mathcal{B}(x, r)$ and let $\Gamma \subset \mathcal{S}$ be a geodesic connecting x and x' . Suppose that Γ contains some vertices and denote them x_1, x_2, \dots, x_k , in order in which they lie on Γ , going from x to x' . Let $x_0 = x$ and $x_{k+1} = x'$. If there is a vertex closer to x_0 than $r_0/2$ and it is not x_1 then we let y_0 be the point at the distance $r_0/2$ from that vertex, between x_0 and x_1 . For every x_j , $j \geq 1$, we let $y_j \in \Gamma$ be the point $r_0/2$ away from x_j , between x_j and x_{j+1} .

Let z_j , $j = 1, \dots, m_1$, be the sequence of all points x_j and y_j , in the order in which they appear on Γ from x to x' , including x and x' . Note that $\sum_{1 \leq j \leq m_1-1} \mathbf{d}(z_j, z_{j+1}) = r$. By the strong Markov property applied at the hitting times of z_j 's, the probability that X starting from x will hit x' in t seconds or less is bounded below by $\prod_{j=1}^{m_1-1} p_j$, where p_j is the probability that X starting from z_j will hit z_{j+1} in t_j seconds or less, and $t_j = t\mathbf{d}(z_j, z_{j+1})/r$.

If z_j is a vertex or x_0 then, by (3.7), if $t_j \leq \mathbf{d}(z_j, z_{j+1})^2$,

$$p_j \geq \frac{1}{m_0} \sqrt{\frac{t_j}{2\pi}} \frac{1}{\mathbf{d}(z_j, z_{j+1})} \exp\left(-\frac{\mathbf{d}(z_j, z_{j+1})^2}{2t_j}\right). \quad (3.8)$$

If $t_j \geq \mathbf{d}(z_j, z_{j+1})^2$ then

$$p_j \geq c_1/m_0 \geq (c_1/m_0) \exp\left(-\frac{\mathbf{d}(z_j, z_{j+1})^2}{2t_j}\right). \quad (3.9)$$

For other z_j 's we use (3.6) to obtain, for $t_j \leq \mathbf{d}(z_j, z_{j+1})^2$,

$$p_j \geq \sqrt{\frac{t_j}{2\pi}} \frac{1}{2\mathbf{d}(z_j, z_{j+1})} \exp\left(-\frac{\mathbf{d}(z_j, z_{j+1})^2}{2t_j}\right), \quad (3.10)$$

and for $t_j \geq \mathbf{d}(z_j, z_{j+1})^2$,

$$p_j \geq c_3 \geq c_3 \exp\left(-\frac{\mathbf{d}(z_j, z_{j+1})^2}{2t_j}\right). \quad (3.11)$$

The product of exponential factors on the right hand sides of (3.8)-(3.11) is equal to

$$\prod_{j=1}^{m_1-1} \exp\left(-\frac{\mathbf{d}(z_j, z_{j+1})^2}{2t_j}\right) = \exp\left(-\frac{r^2}{2t}\right). \quad (3.12)$$

If $t_j \geq \mathbf{d}(z_j, z_{j+1})^2$ then the non-exponential factor in (3.9) is c_1/m_0 and it is c_3 in (3.11). The non-exponential factors in (3.8) and (3.10) are bounded below by

$$\frac{1}{m_0} \sqrt{\frac{t_j}{2\pi}} \frac{1}{2\mathbf{d}(z_j, z_{j+1})} = \frac{1}{m_0} \sqrt{\frac{t\mathbf{d}(z_j, z_{j+1})/r}{2\pi}} \frac{1}{2\mathbf{d}(z_j, z_{j+1})} \geq \frac{1}{m_0} \sqrt{\frac{t}{2\pi}} \frac{1}{2r}.$$

We conclude that the product of non-exponential factors in (3.8)-(3.11) is bounded below by

$$\left(\frac{c_4}{m_0} \sqrt{\frac{t}{2\pi}} \frac{1}{2r}\right)^{m_1-1} \geq \left(\frac{c_4}{m_0} \sqrt{\frac{t}{2\pi}} \frac{1}{2r}\right)^{\lceil r/r_0 \rceil}.$$

This combined with (3.12) gives the lower bound in the lemma.

(ii) Next we will prove the upper bound. Let $\{\Gamma_j\}$ be the family of all Jordan arcs in \mathcal{S} linking x with $\partial\mathcal{B}(x, r)$. The number of edges in $\mathcal{B}(x, r)$ is bounded by $m_2 = m_0^{\lceil r/r_0 \rceil}$ so the number of Γ_j 's is bounded by $m_3 = m_2!$. The length of any Γ_j is at least r .

Consider some Γ_k . We will define a process R_t^k that measures the distance from X_t to x along Γ_k , in a sense. We will “erase” excursions away from Γ_k and loops as follows. For $t > 0$, let $\ell(t) = \sup\{s \leq t : X_s \in \Gamma_k\}$. Let $\mathcal{V}_k = \mathcal{V} \setminus \Gamma_k$, i.e., \mathcal{V}_k is the set of vertices that do not belong to Γ_k . Let $T_1 = 0$,

$$S_j = \inf\{t \geq T_j : X_t \in \mathcal{V}_k\}, \quad j \geq 1, \\ T_{j+1} = \inf\{t \geq S_j : X_t = X_{\ell(S_j)}\}, \quad j \geq 1.$$

If $t \in [S_j, T_{j+1}]$ for some $j \geq 1$, we let R_t^k be the distance from $X_{\ell(S_j)}$ to x along Γ_k . For other t , we let R_t^k be the distance from $X_{\ell(t)}$ to x along Γ_k . The process R_t^k is a time-change of reflected Brownian motion, that is, it is reflected Brownian motion “frozen” on time intervals when X is outside Γ_k (and some other intervals). Hence, the probability that R_t^k reaches r in t seconds or less is less than the right hand side of (3.5). Note that one of the processes R_t^k must be at the level r at the time when X hits $\partial\mathcal{B}(x, r)$. Hence, an upper bound on the probability in the statement of the lemma is the product of the right hand side of (3.5) and m_3 . \square

Example 3.5. Suppose that the graph \mathcal{S} is compact and has the following structure. For some $x \in \mathcal{S}$, the set $\mathcal{S} \setminus \{x\}$ is disconnected and consists of a finite number of disjoint finite

trees $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_k$, and a graph \mathcal{U} (not necessarily a tree). We say that a vertex of a tree is a leaf if it has degree 1. Assume that for some $r_1 > r_2 > 0$ and every leaf $y \neq x$ of any tree \mathcal{T}_j we have $\mathbf{d}(x, y) \geq r_1$, and for any $z \in \mathcal{U}$, $\mathbf{d}(x, z) \leq r_2$ (see, for example, Fig. 3.3). Suppose that (X, Y) is a coupling of Brownian motions on \mathcal{S} . We will show that (X, Y) is not a shy coupling.

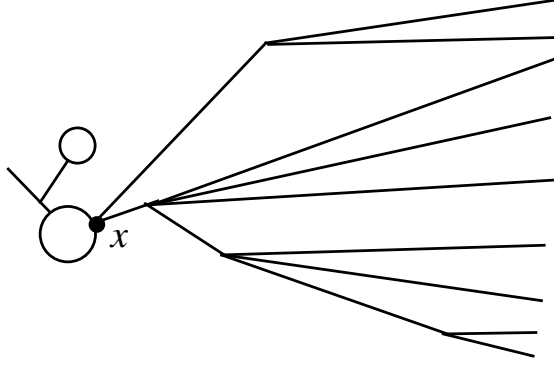


Figure 3.3.

(i) Let $\tilde{\mathcal{T}} = \bigcup_j \mathcal{T}_j \setminus \{x\}$. First, we will show that there exist $p_1 > 0$ and a stopping time $T_1 < \infty$ such that with probability greater than p_1 , either $X_{T_1} = x$ and $Y_{T_1} \in \tilde{\mathcal{T}}$, or $Y_{T_1} = x$ and $X_{T_1} \in \tilde{\mathcal{T}}$. Let $r_3 \in (r_2, r_1)$ and $\mathcal{W} = \{y \in \tilde{\mathcal{T}} : \mathbf{d}(y, x) \geq r_3\}$. It is easy to see that X is recurrent so $T_2 \stackrel{\text{df}}{=} \inf\{t \geq 0 : X_t \in \mathcal{W}\} < \infty$ a.s. Suppose first that $Y_{T_2} \in \tilde{\mathcal{T}} \cup \{x\}$ and let $T_3 = \inf\{t \geq T_2 : X_t = x \text{ or } Y_t = x\}$. Then $T_1 = T_3$ has the properties stated above.

Next suppose that $Y_{T_2} \in \mathcal{U}$ and let $T_4 = \inf\{t \geq T_2 : X_t = x\}$ and $T_5 = \inf\{t \geq T_2 : Y_t = x\}$. By Lemma 3.4, for some $p_2, t_0 > 0$ and all $y \in \mathcal{U}$ and $z \in \mathcal{W}$,

$$\mathbf{P}(T_5 < T_2 + t_0 \mid X_{T_2} = z, Y_{T_2} = y) > \mathbf{P}(T_4 < T_2 + t_0 \mid X_{T_2} = z, Y_{T_2} = y) + p_2.$$

Hence, $\mathbf{P}(T_5 < T_4 \mid X_{T_2} = z, Y_{T_2} = y) > p_2$, and it follows that we can take $T_1 = T_5$ on the event $\{T_5 < T_4\}$. This completes the proof of our claim, with $p_1 = p_2$.

(ii) Recall that the length of any edge is bounded below by $r_0 > 0$. Fix an arbitrarily small $\varepsilon \in (0, r_0/3)$. We will show in the remaining part of the proof that X and Y come within ε distance to each other in finite time almost surely, which will then imply that (X, Y) is not a shy coupling.

The rest of the proof is based on an inductive argument. We will now formulate and prove the inductive step.

Suppose that for some $x_0 \in \mathcal{S}$, $\mathcal{S} \setminus \{x_0\} = \mathcal{U}_1 \cup \mathcal{U}_2$, where \mathcal{U}_1 and \mathcal{U}_2 are disjoint and \mathcal{U}_1 is a finite union of finite trees \mathcal{W}_j , $j = 1, 2, \dots, k$. Assume that $Y_0 = x_0$ and $X_0 \in \mathcal{U}_1$ (the argument is analogous if the roles of X and Y are interchanged). Suppose without loss of generality that $X_0 \in \mathcal{W}_1$. Let $x_1 \neq x_0$ be the vertex of \mathcal{W}_1 closest to x_0 , and $\mathcal{W}_1 \setminus \{x_1\} = \mathcal{W}_2 \cup \mathcal{W}_3$, where \mathcal{W}_2 and \mathcal{W}_3 are disjoint, and \mathcal{W}_3 is the edge joining x_0 and x_1 . We will first assume that $\mathcal{W}_2 \neq \emptyset$. We will show that for some $p_3 > 0$ and some stopping time $T_6 < \infty$, with probability greater than p_3 , we either have $\mathbf{d}(X_{T_6}, Y_{T_6}) \leq \varepsilon$ or $Y_{T_6} = x_1$ and $X_{T_6} \in \mathcal{W}_2$.

If $\mathbf{d}(X_0, Y_0) \leq \varepsilon$ then we can take $T_6 = 0$.

Assume that $\mathbf{d}(X_0, Y_0) > \varepsilon$. Let $x_2 \in \mathcal{W}_3$ be the point with $\mathbf{d}(x_0, x_2) = \varepsilon/3$. Note that $\mathbf{d}(x_2, Y_0) \leq (1/2)\mathbf{d}(x_2, X_0)$. Let $T_7 = \inf\{t \geq 0 : X_t = x_2\}$ and $T_8 = \inf\{t \geq 0 : Y_t = x_2\}$. By Lemma 3.4, for some $s > 0$,

$$\mathbf{P}(T_8 < s) > \mathbf{P}(T_7 < s).$$

Hence, with probability $p_3 > 0$, $T_8 < T_7$ and either X and Y have met by the time T_8 , or X is on the opposite side of Y_{T_8} in \mathcal{S} than x_0 . Let $T_9 = \inf\{t \geq T_8 : Y_t = x_1\}$ and $T_{10} = \inf\{t \geq T_8 : Y_t = x_0\}$. Since \mathcal{S} contains only a finite number of finite trees \mathcal{T}_k , there is an upper bound on the edge length in any tree \mathcal{T}_k , say, $\rho < \infty$. This and the fact that $\mathbf{d}(x_2, x_0) = \varepsilon/3$ imply that $\mathbf{P}(T_9 < T_{10}) \geq p_4$ for some $p_4 > 0$ that may depend on ε . If the events $\{T_8 < T_7\}$ and $\{T_9 < T_{10}\}$ hold then either X and Y have met by the time T_9 or $Y_{T_9} = x_1$ and $X_{T_9} \in \mathcal{W}_2$.

We note that if $\mathcal{W}_2 = \emptyset$ (i.e., \mathcal{W}_1 is a single edge) then the same argument proves that for some $p_3 > 0$ and some stopping time $T_6 < \infty$, we have $\mathbf{d}(X_{T_6}, Y_{T_6}) \leq \varepsilon$ with probability greater than $p_3 p_4$.

(iii) Let us rephrase the claim proved in step (ii). We have shown that for some $p_5 \stackrel{\text{df}}{=} p_3 p_4 > 0$ and some stopping time $T_6 < \infty$, with probability greater than p_5 , we either have $\mathbf{d}(X_{T_6}, Y_{T_6}) \leq \varepsilon$ or Y_{T_6} and X_{T_6} satisfy the same assumptions as Y_0 and X_0 , but relative to graphs $\tilde{\mathcal{U}}_1 \stackrel{\text{df}}{=} \mathcal{W}_2$ and $\tilde{\mathcal{U}}_2 \stackrel{\text{df}}{=} \mathcal{S} \setminus (\{x_2\} \cup \mathcal{W}_2)$ in place of \mathcal{U}_1 and \mathcal{U}_2 . Recall the claim proved in part (i) of the proof and the final remark in step (ii). Note that $\tilde{\mathcal{U}}_1$ has at least one edge less than \mathcal{U}_1 so by induction, we can repeat the inductive step (ii) a finite number of times and show that with a probability $p_6 > 0$, X and Y come within ε of each other before some time $t_1 < \infty$. It is easy to check that p_6 and t_1 can be chosen so that they do not depend on the starting points of X and Y . The Markov property and induction can be used to show that X and Y have to come within ε of each other by the

time jt_1 with probability greater than $1 - (1 - p_6)^j$. We let $j \rightarrow \infty$ to see that X and Y come within ε of each other at some finite time with probability one. Since $\varepsilon \in (0, r_0/3)$ is arbitrary, the coupling is not shy. \square

Example 3.6. Suppose that \mathcal{S} is a tree with the property that it has a “backbone” that is topologically a line, with a finite or countable number of finite trees attached to it. See Fig. 3.4 for an example. Recall that we have assumed that each edge has length at least $r_0 > 0$.

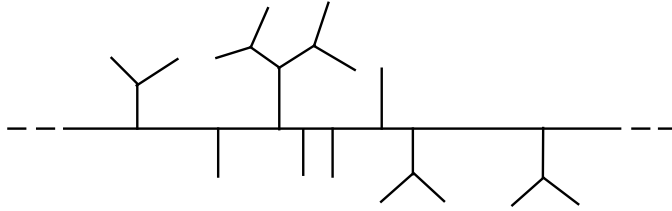


Figure 3.4.

Let $\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots$ be the sequence of points along the “backbone” \mathcal{U} where the side trees are attached (the sequence can be finite or it can extend to infinity in one or two directions). If the sequence extends to infinity in both directions and the graph is invariant under a non-constant shift, then, according to Example 3.3, there exists a shy coupling.

Assume that

- (i) the diameters of the side trees are uniformly bounded and
- (ii) $\{x_k\}$ does not extend to infinity in both directions, or $\{x_k\}$ extends to infinity in both directions but the family $\{x_k\}$ is not shift-invariant, i.e., for every $c \neq 0$, there exists x_j such that $x_j + c \neq x_k$ for all k .

We will show that under these assumptions there is no shy coupling.

Let Q be the “projection” of \mathcal{S} on \mathcal{U} , i.e., $Q(x) = x$ for $x \in \mathcal{U}$, and $Q(x) = x_k$ if x belongs to a tree which is attached to \mathcal{U} at x_k . We will identify the “backbone” \mathcal{U} with the real line so that we can think of $Q(X)$ and $Q(Y)$ as real-valued processes. Let X and Y be a coupling of Brownian motions on \mathcal{S} . If X makes an excursion into a side tree then $Q(X)$ remains constant on the excursion interval (including the endpoints). Hence, $Q(X)$ and $Q(Y)$ are continuous processes. It is easy to see that they are local martingales. Informally speaking, they are Brownian motions frozen on some random intervals. The

process $Z_t = Q(X_t) - Q(Y_t)$ is also a continuous local martingale. Suppose without loss of generality that $Z_0 > 0$ and let

$$T_0 = \inf\{t \geq 0 : Z_t = 0\}.$$

Then $Z_{T_0 \wedge t}$ is a non-negative local martingale and so it must have an almost surely finite limit Z_∞ on $\{T_0 = \infty\}$.

We will show first that $\mathbf{P}(T_0 = \infty) = \mathbf{P}(T_0 = \infty \text{ and } Z_\infty = 0)$, in other words, $\mathbf{P}(T_0 = \infty \text{ and } Z_\infty > 0) = 0$. Let B be a Brownian motion on \mathbb{R} with $B_0 = 0$. Consider a small $\varepsilon > 0$ and $t_0 > 0$, let $\delta \in (0, \varepsilon)$ be such that

$$\mathbf{P}\left(\sup_{t \in [0, t_0]} |B_t| < \delta\right) < p_1/2, \quad \text{where } p_1 := \mathbf{P}\left(\sup_{t \in [0, t_0]} |B_t| < \varepsilon\right).$$

Let T_1 be the first time when all of the following conditions hold: $\mathbf{d}(X_{T_1}, \mathcal{U}) \geq \varepsilon$, the distance from X_{T_1} to any vertex of \mathcal{S} is greater than ε , $Y_{T_1} \in \mathcal{U}$, and $\inf_k \mathbf{d}(Y_{T_1}, x_k) \geq \delta$ (the argument is analogous if the roles of X and Y are interchanged). If $T_1 < \infty$ then with probability $p_1/2$ or greater, X will stay on the same side tree over the interval $[T_1, T_1 + t_0]$, while Y will move away from Y_{T_1} by more than δ units over the same time interval. Hence with probability $p_1/2$ or greater, Z_t will have an oscillation of size at least δ over the interval $[T_1, T_1 + t_0]$. We proceed by induction. If $T_k < \infty$ then we define $T_{k+1} = T_1 \circ \theta_{T_k + t_0} + T_k + t_0$, where θ is the usual Markovian shift operator. Then with probability greater than $p_1/2$, Z_t has an oscillation of size at least δ over the interval $[T_{k+1}, T_{k+1} + t_0]$, independent of whether that happened over any interval $[T_j, T_j + t_0]$, $j \leq k$. Hence, with probability one, either $T_k = \infty$ for some k or Z_t has an infinite number of oscillations of size δ over disjoint intervals of length t_0 , and, therefore in the latter case, Z_t does not have a limit as $t \rightarrow \infty$. Applying the above argument to a decreasing sequence of $\{\varepsilon_n, n \geq 1\}$ and a decreasing sequence of $\{\delta_n, n \geq 1\}$ both tending to zero and after deleting a null set from Ω , we may and do assume that for every $\omega \in \Omega$ and for every ε_n , there is some $N > 1$ such that for every $j \geq N$, with ε_n and δ_j in place of ε and δ above, either $T_k(\omega) = \infty$ for some k or $Z_t(\omega)$ does not have a limit as $t \rightarrow \infty$. The processes X and Y are recurrent because the one-dimensional Brownian motion is. Hence, for every x_j , each one of them will enter the side tree attached to \mathcal{U} at x_j infinitely often. After deleting a null set from Ω , we may and do assume that the aforementioned property holds for every $\omega \in \Omega$.

For $\omega \in \{T_0 = \infty \text{ and } Z_\infty > 0\}$, let

$$c(\omega) = \lim_{t \rightarrow \infty} Z_t(\omega) > 0.$$

We choose an x_j , relative to $c(\omega)$, as follows. If $\{x_n\}$ does not extend to $-\infty$ (∞) then we let x_j be the leftmost (rightmost, resp.) point of the sequence. Otherwise we fix an x_j with the property that $x_j + c(\omega) \neq x_k$ for all k (such an x_j exists by assumption). Note that both $X(\omega)$ and $Y(\omega)$ enter the side tree attached to \mathcal{U} at x_j infinitely often. Hence one can find some $\varepsilon > 0$ from $\{\varepsilon_n, n \geq 1\}$ and $\delta > 0$ from $\{\delta_n, n \geq 1\}$, and an increasing sequence of random times $\{S_k, k \geq 1\}$ with $\lim_{k \rightarrow \infty} S_k = \infty$ such that all of the following hold. One has $\mathbf{d}(X_{S_k}, \mathcal{U}) \geq \varepsilon$, the distance from X_{S_k} to any vertex of \mathcal{S} is greater than ε , $Y_{S_k} \in \mathcal{U}$, and $\inf_n \mathbf{d}(Y_{S_k}, x_n) \geq \delta$ for every $k \geq 1$ (or the statement will hold with the roles of X and Y interchanged).

Hence, all stopping times $\{T_k, k \geq 1\}$ defined in the proceeding paragraph are finite. We have shown that this event implies that $Z_\infty(\omega)$ does not exist. This contradiction proves that $\mathbf{P}(T_0 = \infty \text{ and } Z_\infty > 0) = 0$ and therefore $\mathbf{P}(T_0 = \infty) = \mathbf{P}(T_0 = \infty \text{ and } Z_\infty = 0)$.

Recall that we have assumed that all the edges have length at least $r_0 > 0$. So on $\{T_0 = \infty \text{ and } Z_\infty = 0\}$, by the recurrence of the one-dimensional Brownian motion, we have $\liminf_{t \rightarrow \infty} \mathbf{d}(X_t, Y_t) = 0$. We now only need to exam $\omega \in \{T_0 < \infty\}$ and to prove $X(\omega)$ and $Y(\omega)$ will come arbitrarily close to each other.

Consider any $\varepsilon \in (0, r_0/4)$, where $r_0 > 0$ is a lower bound for the length of any edge in \mathcal{S} . We want to show that with probability one, there exists t such that $\mathbf{d}(X_t, Y_t) \leq \varepsilon$. We have already proved that $\liminf_{t \rightarrow \infty} \mathbf{d}(X_t, Y_t) = 0$ on $\{T_0 = \infty\}$. On $\{T_0 < \infty\}$, at time T_0 , either $X_{T_0} = Y_{T_0}$ or one of processes $\{X_{T_0}, Y_{T_0}\}$ is at some x_k and the other process is in a side tree \mathcal{T} attached to \mathcal{U} at x_k . Without loss of generality, assume that $X_{T_0} = x_k$ and $Y_{T_0} \in \mathcal{T}$. If $\mathbf{d}(X_{T_0}, Y_{T_0}) \leq \varepsilon$ then we are done. Suppose that $\mathbf{d}(X_{T_0}, Y_{T_0}) > \varepsilon$. Let z_0 be the point at the edge e of \mathcal{T} that is attached to \mathcal{U} with $\mathbf{d}(z_0, x_k) = \mathbf{d}(z_0, \mathcal{U}) = \varepsilon/4$. Let S_1 be the first time after T_0 when $X_t = z_0$. By Lemma 3.4, with probability $p_2 > 0$, X_t reaches z_0 after T_0 before Y_t gets there. If this event occurs, both X_{S_1} and Y_{S_1} will have distance at least $\varepsilon/4$ away from \mathcal{U} . Let R_1 be the first time after S_1 when both X_t and Y_t are outside \mathcal{T} . An argument analogous to that in parts (ii) and (iii) of Example 3.5 shows that with probability $p_3 > 0$, the processes X and Y will meet during the time interval $[S_1, R_1]$. In other words, conditioning on $\{T_0 < \infty\}$ and $\mathbf{d}(X_{T_0}, Y_{T_0}) > \varepsilon$, with probability at least $p_4 \stackrel{\text{df}}{=} p_2 p_3 > 0$, the processes X and Y will meet between times T_0 and the first time R_1 when they are both outside \mathcal{T} . We define for $k \geq 2$,

$$S_k = S_1 \circ \theta_{R_{k-1}} + R_{k-1} \quad \text{and} \quad R_k = R_1 \circ \theta_{S_k} + S_k.$$

By the strong Markov property of (X, Y) ,

$$\mathbf{P} \left(\inf_{t \in [0, R_k)} \mathbf{d}(X_t, Y_t) \geq \varepsilon \right) \leq (1 - p_4)^k.$$

Letting $k \rightarrow \infty$, we get

$$\mathbf{P} \left(\inf_{t \in [0, \infty)} \mathbf{d}(X_t, Y_t) \geq \varepsilon \right) \leq 0$$

for every $\varepsilon > 0$ and thus (X, Y) is not a shy coupling. \square

Example 3.7. Suppose that \mathcal{S} is composed of a loop \mathcal{U} with a finite number of finite trees attached to it at points x_k , and the family $\{x_k\}$ is not rotation invariant in the following sense. We can assume without loss of generality that \mathcal{U} is isometric to the unit circle. For every $c \neq 0$, there exists x_j such that $x_j e^{ic} \neq x_k$ for all k . See Fig. 3.5 for an example. We will show that in this case there is no shy coupling.

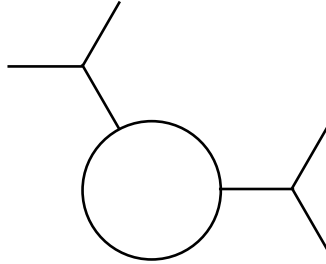


Figure 3.5.

Our argument will be very similar to that in Example 3.6. Recall the “projection” Q from the previous example. We have $Q(x) = x$ for $x \in \mathcal{U}$ and $Q(x) = x_k$ if x belongs to a tree that is attached to \mathcal{U} at x_k . Hence, $Q(X_t)$ may be regarded as a continuous process on the unit circle. We now choose a (random) continuous function $\Theta_X : [0, \infty) \rightarrow \mathbb{R}$ so that $Q(X_t) = e^{i\Theta_X(t)}$ for all $t \geq 0$, in the complex notation. We define Θ_Y in an analogous way. Note that Θ_X and Θ_Y are martingales. Therefore, $Z_t \stackrel{\text{df}}{=} \Theta_X(t) - \Theta_Y(t)$ is also a martingale. We can now repeat the argument from Example 3.6 to show that there does not exist a shy coupling. \square

Example 3.8. Examples 3.3 and 3.5-3.7 may appear to suggest that if a graph has a vertex with degree 1 then a shy coupling exists only if there exists an isometry of $I : \mathcal{S} \rightarrow \mathcal{S}$ with no fixed points. We will show that this is not the case. Our example is illustrated in

Fig. 3.6. In this case, every isometry $I : \mathcal{S} \rightarrow \mathcal{S}$ has a fixed point. Nevertheless, we will show that there is a shy coupling in \mathcal{S} .

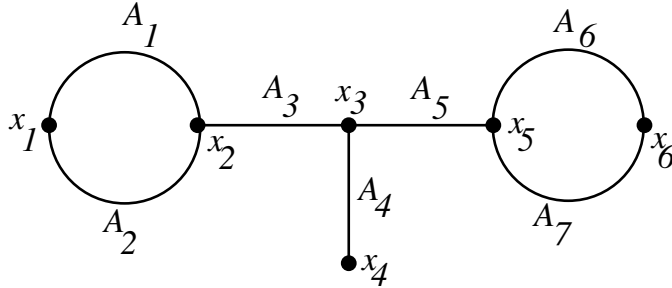


Figure 3.6.

We will describe below the transition mechanism for (X, Y) on some random intervals of time. We will assume that the transition probabilities of (Y, X) are the same as those of (X, Y) . Hence, there is no need to describe cases symmetric to those discussed below, in the sense that the initial positions of X and Y are interchanged.

Suppose that all edges A_1, A_2, \dots, A_7 have the same length, say 1. We assume that $X_0 = x_2$ and $Y_0 = x_3$.

(i) Suppose that for some stopping time T_1 we have $X_{T_1} = x_2$ and $Y_{T_1} = x_3$. Then we let $T_2 = \inf\{t \geq T_1 : X_t \notin (A_1 \cup A_2 \cup A_3) \setminus \{x_1\}\}$. Let $I : (A_1 \cup A_2 \cup A_3) \setminus \{x_1, x_3\} \rightarrow (A_3 \cup A_4 \cup A_5) \setminus \{x_2, x_4, x_5\}$ be the one-to-one isometry satisfying $I(x_2) = x_3$, $I(A_1) = A_3$, $I(A_2) = A_4$ and $I(A_3) = A_5$. We let $Y_t = I(X_t)$ for $t \in [T_1, T_2]$. Note that at the stopping time T_2 , we have one of the following configurations of the two particles: $(X_{T_2}, Y_{T_2}) = (x_1, x_4)$, or $(X_{T_2}, Y_{T_2}) = (x_1, x_2)$, or $(X_{T_2}, Y_{T_2}) = (x_3, x_5)$.

(ii) Suppose that for some stopping time T_3 we have $X_{T_3} = x_1$ and $Y_{T_3} = x_4$. Let $\{X_t, t \in [T_3, T_4]\}$ be Brownian motion on \mathcal{S} independent of the past with $X_{T_3} = x_3$, where $T_4 = \inf\{t \geq T_3 : X_t = x_2\}$. For $t \in [T_3, T_4]$, we let $Y_t \in A_4$, with $\mathbf{d}(Y_t, x_4) = \mathbf{d}(X_t, x_1)$. Note that $X_{T_4} = x_2$ and $Y_{T_4} = x_3$.

(iii) Suppose that for some stopping time T_5 we have $X_{T_5} = x_1$ and $Y_{T_5} = x_2$. Let $\{Y_t, t \in [T_5, T_6]\}$ be Brownian motion on \mathcal{S} independent of the past with $Y_{T_5} = x_2$, where $T_6 = \inf\{t \geq T_5 : Y_t = x_1 \text{ or } x_3\}$. We label excursions of Y from x_2 that stay in A_3 with marks “1” or “2”, with equal probabilities, in such a way that the label of any excursion is independent of all other labels. Then we let X_t be defined for $t \in [T_5, T_6]$ by $\mathbf{d}(X_t, x_1) = \mathbf{d}(Y_t, x_2)$ and the following conditions. If $Y_t \in A_1$ then $X_t \in A_2$, if $Y_t \in A_2$ then $X_t \in A_1$, if $Y_t \in A_3$ and t belongs to an excursion marked “1” then $X_t \in A_1$, and

if $Y_t \in A_3$ and t belongs to an excursion marked “2” then $X_t \in A_2$. At time T_6 we have $(X_{T_6}, Y_{T_6}) = (x_2, x_1)$ or $(X_{T_6}, Y_{T_6}) = (x_2, x_3)$.

Note that \mathcal{S} is symmetric with respect to the line containing A_4 . If for some stopping time T_7 we have $(X_{T_7}, Y_{T_7}) = (x_3, x_5)$, or $(X_{T_7}, Y_{T_7}) = (x_4, x_6)$, or $(X_{T_7}, Y_{T_7}) = (x_5, x_6)$, or one of these conditions is satisfied with the roles of X and Y interchanged, then we define the coupling on an appropriate random interval in a way analogous to that in (i)-(iii), using the symmetry of \mathcal{S} .

The above definitions for the “local” behavior of the coupling and the strong Markov property can now be used to define a process (X_t, Y_t) for all $t \geq 0$. It is easy to see that the stopping times analogous to T_1, T_3 and T_5 will not have a finite point of accumulation. It is also easy to check that almost surely $\mathbf{d}(X_t, Y_t) = 1$ for every $t > 0$. \square

4. Reflected Brownian motion in Euclidean domains.

This section is the closest in spirit to the papers and problems which inspired the present research project. Suppose that $D \subset \mathbb{R}^d$ is a bounded connected open set which is either convex or has a C^2 boundary. We will consider couplings (X, Y) of reflected Brownian motions in D , defined as follows. Let $\mathbf{n}(x)$ denote the unit inward normal vector at $x \in \partial D$. Let B and W be standard planar Brownian motions with $B_0 = W_0 = 0$ defined on the same probability space and consider the following Skorohod equations,

$$X_t = x_0 + B_t + \int_0^t \mathbf{n}(X_s) dL_s^X, \quad (4.1)$$

$$Y_t = y_0 + W_t + \int_0^t \mathbf{n}(Y_s) dL_s^Y. \quad (4.2)$$

Here L^X is the local time of X on ∂D , i.e., a non-decreasing continuous process which does not increase when X is in D : $\int_0^\infty \mathbf{1}_D(X_t) dL_t^X = 0$, a.s. Equation (4.1) has a unique pathwise solution (X, L^X) such that $X_t \in \bar{D}$ for all $t \geq 0$ (see [Ta] when D is convex domain and [LS] when D is C^2). The “reflected Brownian motion” X is a strong Markov process. We point out that B is uniquely determined by X , and vice versa. The same remarks apply to (4.2), so, as a pair, (X, Y) is also strong Markov.

For a continuous semimartingale M , the symbol $\langle Z \rangle$ will stand for its quadratic variation process. When $M = (M^1, \dots, M^d)$ and $N = (Z^1, \dots, Z^d)$ are two continuous \mathbb{R}^d -valued semimartingales we will use $\langle M, N \rangle$ to denote $\sum_{i,j=1}^d \langle M^i, N^j \rangle$. Note that the matrix-valued process $(\langle M^i, N^j \rangle)_{1 \leq i,j \leq d}$ is non-negative definite and so $t \mapsto \langle M, M \rangle_t$ is always non-decreasing. For $a, b \in \mathbb{R}^d$, we use $a \cdot b$ to denote the inner product between a

and b . We will use $\mathbf{d}(x, y)$ and $|x - y|$ interchangeably for the Euclidean distance between $x, y \in \mathbb{R}^d$.

Theorem 4.1. *Assume that $D \subset \mathbb{R}^d$ is a bounded convex domain. Let X and Y be two reflecting Brownian motion on D given by (4.1)-(4.2).*

(i) *Suppose that there is a strictly increasing function φ with $\varphi(0) = 0$ such that*

$$d\langle |X - Y|^2 \rangle_t \geq \varphi(|X_t - Y_t|) dt \quad \text{for } t < \sigma_0,$$

where $\sigma_0 := \inf\{t > 0 : X_t = Y_t\}$. Then (X, Y) is not a shy coupling.

(ii) *Suppose D is strictly convex. Assume that $\langle X - Y, X - Y \rangle_t$ (this is the same as $\langle B - W, B - W \rangle_t$) has a sublinear growth rate as $t \rightarrow \infty$, that is,*

$$\lim_{t \rightarrow \infty} \langle X - Y, X - Y \rangle_t / t = 0 \quad \text{almost surely.}$$

Then (X, Y) is not a shy coupling.

Proof. Note that

$$X_t - Y_t = X_0 - Y_0 + (B - W) + \int_0^t \mathbf{n}(X_s) dL_s^X - \int_0^t \mathbf{n}(Y_s) dL_s^Y$$

is a semimartingale. Define $R_t := |X_t - Y_t|^2$. By Ito's formula,

$$\begin{aligned} dR_t &= 2(X_t - Y_t) \cdot d(X_t - Y_t) + d\langle X - Y, X - Y \rangle_t \\ &= 2(X_t - Y_t) \cdot d(B_t - W_t) - 2(Y_t - X_t) \cdot \mathbf{n}(X_t) dL_t^X - 2(X_t - Y_t) \cdot \mathbf{n}(Y_t) dL_t^Y \\ &\quad + d\langle X - Y, X - Y \rangle_t. \end{aligned} \tag{4.3}$$

(i) Let $a > 0$ be a constant whose value will be chosen in a moment and $f(r) = -r^{-a}$ for $r > 0$. Then $f'(r) = ar^{-a-1} > 0$ and $f''(r) = (-a - 1)ar^{-a-2} < 0$ for $r > 0$. Define $U_t := f(R_t) = f(|X_t - Y_t|^2)$. By Ito's formula, we have

$$dU_t = f'(R_t) dR_t + \frac{1}{2} f''(R_t) d\langle R \rangle_t = dM_t + dV_t,$$

where

$$dM_t = 2aR_t^{-a-1} 2(X_t - Y_t) \cdot d(B_t - W_t)$$

and

$$\begin{aligned} dV_t &= -2aR_t^{-a-1} ((Y_t - X_t) \cdot \mathbf{n}(X_t) dL_t^X + (X_t - Y_t) \cdot \mathbf{n}(Y_t) dL_t^Y) \\ &\quad + aR_t^{-a-1} d\langle X - Y, X - Y \rangle_t - 2a(a + 1)R_t^{-a-2} d\langle |X - Y|^2 \rangle_t \end{aligned} \tag{4.4}$$

are the local martingale and bounded variation parts, respectively. We claim that for every $\varepsilon > 0$,

$$T_\varepsilon := \inf \{t > 0 : |X_t - Y_t| \leq \varepsilon\}$$

is finite almost sure. Suppose that $\mathbf{P}(T_\varepsilon = \infty) > 0$ for some $\varepsilon > 0$. We will show that this leads to a contradiction.

Since D is a convex domain, for σ -a.e. $x \in \partial D$, $\mathbf{n}(x)$ is well defined and

$$(y - x) \cdot \mathbf{n}(x) \geq 0 \quad \text{for every } y \in \overline{D}.$$

Note that the local time L^X (respectively, L^Y) does not increase when X (respectively Y) is on a subset of ∂D having zero Lebesgue surface measure. Hence, (4.4) yields

$$dV_t \leq aR_t^{-a-1}d\langle X - Y, X - Y \rangle_t - 2a(a+1)R_t^{-a-2}d\langle |X - Y|^2 \rangle_t. \quad (4.5)$$

Note that $d\langle X - Y, X - Y \rangle_t = d\langle B - W, B - W \rangle_t \leq 4dt$. This, (4.5) and the hypothesis in part (i) of this theorem imply that on $\{T_\varepsilon = \infty\}$,

$$\begin{aligned} dV_t &\leq 4aR_t^{-a-1}dt - 2a(a+1)R_t^{-a-2}\varphi(\varepsilon)dt \\ &\leq -2aR_t^{-a-2}((a+1)\varphi(\varepsilon) - 2R_t)dt \\ &\leq -2a\varepsilon^{-a-2}((a+1)\varphi(\varepsilon) - 2\text{diam}(D))dt. \end{aligned}$$

For a fixed $\varepsilon > 0$, we can find $a > 0$ sufficiently large so that for some $\lambda > 0$,

$$dV_t \leq -\lambda dt \quad \text{for every } t > 0 \text{ on } \{T_\varepsilon = \infty\}. \quad (4.6)$$

The continuous local martingale M is a time change of Brownian motion. By the law of iterated logarithm for Brownian sample path, for almost all $\omega \in \{T_\varepsilon = \infty\}$, there is an unbounded increasing sequence $\{t_k, k \geq 1\}$ such that $\sup_{k \geq 1} |M_{t_k}(\omega)| < \infty$. This and (4.6) imply that $U_{t_k}(\omega) = V_0(\omega) + M_{t_k}(\omega) + V_{t_k}(\omega)$ tends to $-\infty$ as $k \rightarrow \infty$ on $\{T_\varepsilon = \infty\}$ a.s. Consequently, $|X_{t_k} - Y_{t_k}|$ goes to 0 as $k \rightarrow \infty$ on $\{T_\varepsilon = \infty\}$ a.s., which is a contradiction. This proves that particles X and Y come arbitrarily close to each other in finite time and, therefore, X and Y is not a shy coupling.

(ii) Now assume that D is bounded and strictly convex and $\langle X - Y, X - Y \rangle_t$ has a sublinear growth as $t \rightarrow \infty$. The strict convexity implies (in fact, it is equivalent to) the following condition. For every small $\varepsilon > 0$, there is a constant $a_\varepsilon > 0$ such that

$$(y - x) \cdot \mathbf{n}(x) \geq a_\varepsilon |x - y| \quad \text{for every } x \in \partial D \text{ and } y \in \overline{D} \text{ with } |x - y| \geq \varepsilon. \quad (4.7)$$

Let σ denote the surface measure on ∂D . Since reflecting Brownian motion in D is a recurrent Feller process, it follows from the Ergodic Theorem that

$$\lim_{t \rightarrow \infty} \frac{L^X}{t} = \frac{\sigma(\partial D)}{2|D|} = \lim_{t \rightarrow \infty} \frac{L^Y}{t} \quad \text{almost surely.}$$

For every $\varepsilon > 0$, define $T_\varepsilon := \inf\{t > 0 : |X_t - Y_t| \leq \varepsilon\}$. On $\{T_\varepsilon = \infty\}$, we have from above and (4.7) that

$$\begin{aligned} & \liminf_{t \rightarrow \infty} \frac{1}{t} \left(-2 \int_0^t (Y_t - X_t) \cdot \mathbf{n}(X_t) dL_t^X - 2 \int_0^t (X_t - Y_t) \cdot \mathbf{n}(Y_t) dL_t^Y + \langle X - Y, X - Y \rangle_t \right) \\ & \leq \liminf_{t \rightarrow \infty} \frac{1}{t} \left(-2\varepsilon a_\varepsilon L_t^X - 2\varepsilon a_\varepsilon L_t^Y + \langle X - Y, X - Y \rangle_t \right) = -\frac{2\varepsilon a_\varepsilon \sigma(\partial D)}{|D|} < 0. \end{aligned} \quad (4.8)$$

On the other hand, $M_t := 2 \int_0^t (X_t - Y_t) \cdot d(B_t - W_t)$ is a continuous martingale and thus is a time-change of one-dimensional Brownian motion. By the law of iterated logarithm for Brownian sample path, for almost all $\omega \in \{T_\varepsilon = \infty\}$, there is an unbounded increasing sequence $\{t_k, k \geq 1\}$ such that $\sup_{k \geq 1} |M_{t_k}(\omega)| < \infty$. This, (4.3) and (4.8) imply that $\lim_{t \rightarrow \infty} R_t = -\infty$ a.s. on $\{T_\varepsilon = \infty\}$. Since $R_t \geq 0$, we conclude that $\mathbf{P}(T_\varepsilon = \infty) = 0$ for every $\varepsilon > 0$ and so (X, Y) is not a shy coupling. \square

In the remainder of this section, we take $d = 2$, but this is only for notational convenience. We will show in the next example that the method of proof of Theorem 4.1, based on the Itô formula, does not extend to arbitrary couplings. The example may have some interest of its own. We will show in Theorem 4.3 below that, in fact, there is no shy coupling of reflecting Brownian motions on any bounded C^1 -smooth strictly convex domain.

Example 4.2. We will show that there exist planar Brownian motions B and W with the property that $\mathbf{d}(B_t, W_t) = \sqrt{2t + \mathbf{d}(B_0, W_0)^2}$ for $t \geq 0$, assuming that $B_0 \neq W_0$. In particular, the distance between the two processes grows in a deterministic way.

Suppose that (B, W) has the above mentioned property with B_0 and W_0 taking values in \overline{D} . Let X and Y be the pathwise solutions of (4.1)-(4.2) but with the above B and W in place of $x_0 + B$ and $y_0 + W$ there. We have

$$d\langle |X_t - Y_t|^2 \rangle = d\langle |B_t - W_t|^2 \rangle = 0,$$

while

$$d\langle X - Y, X - Y \rangle_t = d\langle B - W, B - W \rangle_t = d(|B_t - W_t|^2) = 2t.$$

So V_t in (4.4) becomes

$$\begin{aligned} dV_t &= -2aR_t^{-a-1} \left((Y_t - X_t) \cdot \mathbf{n}(X_t) dL_t^X + (X_t - Y_t) \cdot \mathbf{n}(Y_t) dL_t^Y \right) \\ &\quad + 2aR_t^{-a-1} dt. \end{aligned}$$

Hence the method used in the proof of Theorem 4.1 does not work for this coupling (X, Y) . Moreover, since $|X_t - Y_t|$ grows deterministically when both X_t and Y_t are away from the boundary and decreases when one of them is on the boundary, neither $|X_t - Y_t|$ nor any deterministic monotone function of $|X_t - Y_t|$ is a submartingale or a supermartingale.

We now present the construction of B and W with the properties mentioned above. Let B be a Brownian motion in \mathbb{R}^2 starting from x_0 . For a vector $v = (a, b) \in \mathbb{R}^2$, we use v^\perp to denote its orthogonal vector $(b, -a)$. Let $y_0 \in \mathbb{R}^2$ be a point different from x_0 . Consider the following SDE for W in \mathbb{R}^2 with $W_0 = y_0$:

$$dW_t = \frac{1}{|W_t - B_t|^2} \left(((W_t - B_t) \cdot dB_t) (W_t - B_t) - ((W_t - B_t)^\perp \cdot dB_t) (W_t - B_t)^\perp \right).$$

In words, at any given time $t > 0$, W_t takes a synchronous step with B_t along the direction $W_t - B_t$, while W_t moves in the opposite direction but with the same magnitude as B_t along the perpendicular direction $(W_t - B_t)^\perp$. The above SDE for W has a unique solution up to $\tau := \inf\{t > 0 : W_t = B_t\}$, since the diffusion coefficients are C^∞ up to that time. It can be computed directly that $d(|W_t - B_t|^2) = 2dt$ and consequently $|W_t - B_t|^2 = |x_0 - y_0|^2 + 2t$. So $\tau = \infty$. It is standard to check that $W = (W^1, W^2)$ is a continuous local martingale with $\langle W^i, W^i \rangle_t = t$ for $i = 1, 2$ and $\langle W^1, W^2 \rangle = 0$. Therefore W is a Brownian motion in \mathbb{R}^2 starting from y_0 . \square

We will show next that in a C^1 -smooth strictly convex domain D , every coupling of reflecting Brownian motions on \overline{D} must come arbitrarily close to each other in finite time.

Theorem 4.3. *Suppose that D is a bounded convex planar domain with a C^1 -smooth boundary that does not contain any line segments. Then there does not exist a shy coupling (X, Y) of reflected Brownian motions in D .*

Proof. The idea of the proof is inspired by differential games of pursuit (see [F]). We will show that with positive probability, one of the particles will pursue the other one in such a way that the distance between the two particles decreases either because the diffusion

component of the second process does not move the second particle sufficiently fast or the second particle hits the boundary and is pushed back towards the first one.

Step 1. We will define several constants ε_k in this step. The definitions will be labeled (a), (b), (c), etc. Each of these definitions is really a simple lemma asserting the existence of a constant with stated properties. Since the proofs do not need more than high school geometry, we omit most of the proofs. The constants ε_k are defined relative to each other, but ε_k may depend only on the values of ε_j for $j < k$.

(a) Let $\varepsilon_0 > 0$ be so small that for every $x \in D$ with $\mathbf{d}(x, \partial D) \leq \varepsilon_0$ there exists a unique point in ∂D whose distance from x is minimal.

We make $\varepsilon_0 > 0$ smaller, if necessary, so that the following is true. Consider any point $y \in \partial D$ and let CS_1 be the orthonormal coordinate system such that $y = 0 \in \partial D$ and $\mathbf{n}(0)$ lies on the second axis. Write $\mathbf{n}(x) = (\mathbf{n}_1(x), \mathbf{n}_2(x))$. Then $|\mathbf{n}_1(x)| \leq \mathbf{n}_2(x)/100$ for $x \in \partial D \cap \mathcal{B}(0, \varepsilon_0)$ in CS_1 .

We fix an arbitrary $\varepsilon_1 \in (0, \varepsilon_0]$. It will suffice to prove that for any $x_0, y_0 \in \overline{D}$, if $(X_0, Y_0) = (x_0, y_0)$ then, with probability one, there exists $t < \infty$ such that $\mathbf{d}(X_t, Y_t) \leq \varepsilon_1$.

(b) The angle between two vectors will be denoted $\angle(\cdot, \cdot)$, with the convention that it takes values in $(-\pi, \pi]$. Since D is a bounded and strictly convex domain, there exists $\varepsilon_2 \in (0, \pi/2)$ such that for every $x \in \partial D$ and $y \in \overline{D}$ satisfying $\mathbf{d}(x, y) \geq \varepsilon_1/2$,

$$\angle(\mathbf{n}(x), y - x) \in [-\pi/2 + \varepsilon_2, \pi/2 - \varepsilon_2]. \quad (4.9)$$

(c) Let $\mathcal{L}(x, r)$ be the cone spanned by $\{\mathbf{n}(y), y \in \partial D \cap \mathcal{B}(x, r)\}$. Since ∂D is C^1 -smooth, $\mathcal{L}(x, r)$ is a wedge. Hence, all linear combinations of vectors in $\mathcal{L}(x, r)$ with non-negative coefficients belong to $\mathcal{L}(x, r)$. An easy approximation argument shows that if $X_t \in \mathcal{B}(x, r)$ for all $t \in (s, u)$ then $\int_s^u \mathbf{n}(X_t) dL_t^X \in \mathcal{L}(x, r)$.

We will now choose $\varepsilon_3 \in (0, \varepsilon_1/8)$. Consider vectors \mathbf{v} and \mathbf{w} satisfying the following conditions, relative to x_0, y_0 , and ε_3 .

If $\mathbf{d}(x_0, \partial D) \leq \varepsilon_3$ then $\mathbf{v} \in \mathcal{L}(x_0, 2\varepsilon_3)$ and $|\mathbf{v}| \leq \varepsilon_3$. If $\mathbf{d}(x_0, \partial D) > \varepsilon_3$ then $\mathbf{v} = 0$.

If $\mathbf{d}(y_0, \partial D) \leq \varepsilon_3$ then $\mathbf{w} \in \mathcal{L}(y_0, 2\varepsilon_3)$ and $|\mathbf{w}| \leq \varepsilon_3$. If $\mathbf{d}(y_0, \partial D) > \varepsilon_3$ then $\mathbf{w} = 0$.

We will show that (4.9) implies that we can find sufficiently small $\varepsilon_3 > 0$ so that the following is true. Suppose that $x_0, y_0 \in \overline{D}$ with $\mathbf{d}(x_0, y_0) \geq \varepsilon_1$. Assume that $x_1 \in \mathcal{B}(x_0, 2\varepsilon_3)$ and $y_1 \in \mathcal{B}(y_0, 2\varepsilon_3)$. Then

$$\mathbf{d}(x_1 + \mathbf{v}, y_1 + \mathbf{w}) \leq \mathbf{d}(x_1, y_1). \quad (4.10)$$

To see this, choose some x_2 and y_2 so that the following conditions hold.

If $\mathbf{d}(x_0, \partial D) \leq \varepsilon_3$ then $x_2 \in \partial D \cap \mathcal{B}(x_0, 2\varepsilon_3)$ with $\mathbf{v} = \mathbf{cn}(x_2)$. If $\mathbf{d}(x_0, \partial D) > \varepsilon_3$ then $x_2 = x_0$.

If $\mathbf{d}(y_0, \partial D) \leq \varepsilon_3$ then $y_2 \in \partial D \cap \mathcal{B}(y_0, 2\varepsilon_3)$ with $\mathbf{w} = \mathbf{cn}(y_2)$. If $\mathbf{d}(y_0, \partial D) > \varepsilon_3$ then $y_2 = y_0$.

Since $\varepsilon_3 < \varepsilon_1/8$,

$$|x_2 - y_2| \geq |x_0 - y_0| - |x_0 - x_2| - |y_0 - y_2| \geq \varepsilon_1 - 4\varepsilon_3 \geq \varepsilon_1/2.$$

Thus by (4.9), we have for sufficiently small $\varepsilon_3 > 0$,

$$\begin{aligned} & \mathbf{d}(x_1 + \mathbf{v}, y_1 + \mathbf{w})^2 \\ & \leq |x_1 - y_1|^2 + |\mathbf{v} - \mathbf{w}|^2 + 2(x_1 - y_1) \cdot (\mathbf{v} - \mathbf{w}) \\ & \leq |x_1 - y_1|^2 + |\mathbf{v} - \mathbf{w}|^2 + 2(x_2 - y_2) \cdot (\mathbf{v} - \mathbf{w}) + 2(|x_1 - x_2| + |y_1 - y_2|)|\mathbf{v} - \mathbf{w}| \\ & \leq |x_1 - y_1|^2 + |\mathbf{v} - \mathbf{w}|^2 - 2(\sin \varepsilon_2)|x_2 - y_2|(|\mathbf{v}| + |\mathbf{w}|) + 4(8\varepsilon_3)|\mathbf{v} - \mathbf{w}| \\ & \leq |x_1 - y_1|^2 + (|\mathbf{v}| + |\mathbf{w}|)(6\varepsilon_3 - 2(\sin \varepsilon_2)(\varepsilon_1/2) + 32\varepsilon_3) \\ & = |x_1 - y_1|^2 - (|\mathbf{v}| + |\mathbf{w}|)(\varepsilon_1 \sin \varepsilon_2 - 38\varepsilon_3) \\ & \leq \mathbf{d}(x_1, y_1)^2. \end{aligned} \tag{4.11}$$

(d) Since D is bounded, we can find $\varepsilon_4 > 0$ and $N < \infty$, such that if x_1, x_2, \dots is a sequence of points with $x_1 \in \overline{D}$, $\mathbf{d}(x_k, x_{k-1}) \geq \varepsilon_3/8$ and $|\angle(x_k - x_{k-1}, x_{k+1} - x_k)| \leq \varepsilon_4$ for all k then $x_N \notin \overline{D}$.

(e) We choose $\varepsilon_5, \varepsilon_6 > 0$ so that the following is true. Suppose that x_0, y_0, x_1, y_1 and x_2 satisfy the conditions $\mathbf{d}(x_0, y_0) \geq \varepsilon_1$, $\mathbf{d}(x_1, y_1) \geq \varepsilon_1$ and

$$\begin{aligned} & |\angle(x_1 - y_1, x_0 - y_0)| \leq \varepsilon_5, \\ & \mathbf{d}(x_1, x_0 + (\varepsilon_3/4)(y_0 - x_0)/\mathbf{d}(x_0, y_0)) \leq \varepsilon_6, \\ & \mathbf{d}(x_2, x_1 + (\varepsilon_3/4)(y_1 - x_1)/\mathbf{d}(x_1, y_1)) \leq \varepsilon_6. \end{aligned}$$

Then $|\angle(x_1 - x_0, x_2 - x_1)| \leq \varepsilon_4$.

We make ε_6 smaller, if necessary, so that $\varepsilon_6 < \varepsilon_3/8$.

(f) We make $\varepsilon_6 > 0$ smaller, if necessary so that the following holds. Suppose that x_0, y_0, x_1 and y_1 satisfy the conditions $\mathbf{d}(x_0, y_0) \geq \varepsilon_1$,

$$\mathbf{d}(x_1, x_0 + (\varepsilon_3/4)(y_0 - x_0)/\mathbf{d}(x_0, y_0)) \leq \varepsilon_6, \tag{4.12}$$

$$\mathbf{d}(y_1, y_0 + (\varepsilon_3/4)(y_0 - x_0)/\mathbf{d}(x_0, y_0)) \leq \varepsilon_6. \tag{4.13}$$

Then $|\angle(x_1 - y_1, x_0 - y_0)| \leq \varepsilon_5$.

(g) We can find $\varepsilon_7, \varepsilon_8 > 0$ with the following properties. Suppose that $x_0, y_0 \in \bar{D}$, $x_1, x_2 \in \mathbb{R}^2$, $\mathbf{d}(x_0, y_0) \geq \varepsilon_1$ and the following conditions are satisfied.

If $\mathbf{d}(x_0, \partial D) \leq \varepsilon_3$ then $\mathbf{v} \in \mathcal{L}(x_0, 2\varepsilon_3)$ and $|\mathbf{v}| \leq \varepsilon_3$. If $\mathbf{d}(x_0, \partial D) > \varepsilon_3$ then $\mathbf{v} = 0$.

If $\mathbf{d}(y_0, \partial D) \leq \varepsilon_3$ then $\mathbf{w} \in \mathcal{L}(y_0, 2\varepsilon_3)$ and $|\mathbf{w}| \leq \varepsilon_3$. If $\mathbf{d}(y_0, \partial D) > \varepsilon_3$ then $\mathbf{w} = 0$.

Assume that

$$\mathbf{d}(x_1, x_0 + (\varepsilon_3/4)(y_0 - x_0)/\mathbf{d}(x_0, y_0)) \leq \varepsilon_8, \quad (4.14)$$

$$\mathbf{d}(y_1, y_0 + (\varepsilon_3/4)(y_0 - x_0)/\mathbf{d}(x_0, y_0)) \geq \varepsilon_6/2, \quad (4.15)$$

$$\mathbf{d}(y_1, y_0) \leq \varepsilon_3/4 + \varepsilon_8. \quad (4.16)$$

Then $\mathbf{d}(x_1 + \mathbf{v}, y_1 + \mathbf{w}) \leq \mathbf{d}(x_0, y_0) - \varepsilon_7$.

(h) It is easy to see from (4.11) that we can strengthen (4.10) as follows. We can make $\varepsilon_7 > 0$ smaller, if necessary, so that if $|\mathbf{v}| \geq \varepsilon_6/2$ or $|\mathbf{w}| \geq \varepsilon_6/2$, and the assumptions stated in Step 1(c) hold then

$$\mathbf{d}(x_1 + \mathbf{v}, y_1 + \mathbf{w}) \leq \mathbf{d}(x_1, y_1) - 2\varepsilon_7. \quad (4.17)$$

Step 2. Suppose that $X_0, Y_0 \in \bar{D}$ with $\mathbf{d}(X_0, Y_0) \geq \varepsilon_1$. Consider the following events,

$$F_1(t) = \{\mathbf{d}(X_t, Y_t) \leq \mathbf{d}(X_0, Y_0) - \varepsilon_7\},$$

$$F_2(t) = \{|\angle(X_t - Y_t, X_0 - Y_0)| \leq \varepsilon_5\},$$

$$F_3(t) = \{\mathbf{d}(X_t, X_0 + (\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0)) \leq \varepsilon_6 \wedge \varepsilon_8\},$$

$$F_4(t) = \{\mathbf{d}(X_t, Y_t) \leq \mathbf{d}(X_0, Y_0) + \varepsilon_7/(4N)\},$$

$$F_5(t) = (F_1(t) \cup F_2(t)) \cap F_3(t) \cap F_4(t).$$

We will show in Step 4 that $\mathbf{P}(F_5(t_1)) > p_1$ for some $t_1, p_1 > 0$ that do not depend on X_0 and Y_0 .

Let $\varepsilon_9 = \varepsilon_7/(16N) \wedge \varepsilon_3/8 \wedge \varepsilon_6/8 \wedge \varepsilon_8/5$. Recall that B and W are Brownian motions with $B_0 = W_0 = 0$ driving X and Y in the sense of (4.1)-(4.2) and let

$$A_1(t) = \{B_t \in \mathcal{B}((\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0), \varepsilon_9)\},$$

$$A_2(t) = \left\{ \sup_{s \in [0, t]} |B_s - (s/t)B_t| \leq \varepsilon_9 \right\},$$

$$A_3(t) = \{|W_t| \leq \varepsilon_3/4 + \varepsilon_9\},$$

$$A_4(t) = \left\{ \sup_{s \in [0, t]} |W_s - (s/t)W_t| \leq \varepsilon_9 \right\},$$

$$A_5(t) = A_1(t) \cap A_2(t) \cap A_3(t) \cap A_4(t).$$

We will argue in the rest of this step of the proof that $\mathbf{P}(A_5(t_1)) > p_1$ for some $p_1, t_1 > 0$. In later steps, we will show that $A_5(t) \subset F_5(t)$.

Recall that B is a two-dimensional Brownian motion with $B_0 = 0$ and let $T_r = \inf\{t \geq 0 : |B_t| > r\}$. Note that, by Brownian scaling,

$$\mathbf{P}(T_r < t) = \mathbf{P}\left(\max_{0 \leq s \leq t} |B_s| > r\right) = \mathbf{P}\left(\max_{0 \leq s \leq 1} |B_s| > r/\sqrt{t}\right).$$

By the large deviations principle (see [RY], Ch. VIII, Thm. 2.11),

$$\lim_{t/r^2 \rightarrow 0} (2t/r^2) \log \mathbf{P}(T_r < t \mid B_0 = 0) = -1. \quad (4.18)$$

Let

$$\begin{aligned} r_0 &= \varepsilon_3/4, \\ A_6(t) &= \{T_{r_0} < t\}, \\ A_7 &= \{B_{T_{r_0}} \in \mathcal{B}((\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0), \varepsilon_9/2)\}, \\ A_8(t) &= \left\{ \sup_{T_{r_0} \leq s \leq T_{r_0} + t} |B_s - B_{T_{r_0}}| \leq \varepsilon_9/2 \right\}. \end{aligned}$$

Clearly $A_6(t)$, A_7 and $A_8(t)$ are independent, and $A_6(t) \cap A_7 \cap A_8(t) \subset A_1(t)$. So

$$\mathbf{P}(A_1(t)) \geq \mathbf{P}_6(A_6(t)) \cdot \mathbf{P}(A_7) \cdot \mathbf{P}(A_8(t)).$$

Note that $\lim_{t \rightarrow 0} P(A_8(t)) = 1$ by the strong Markov property applied at T_{r_0} and the event A_7 is independent of t . This together with (4.18) yields

$$\liminf_{t \rightarrow 0} t \log \mathbf{P}(A_1(t)) \geq -r_0^2/2 = -(\varepsilon_3/4)^2/2.$$

We obtain directly from (4.18) that

$$\limsup_{t \rightarrow 0} t \log \mathbf{P}(A_3^c(t)) \leq -(\varepsilon_3/4 + \varepsilon_9)^2/2.$$

This implies that for all sufficiently small $t > 0$ we have

$$\mathbf{P}(A_1(t) \cap A_3(t)) > 0. \quad (4.19)$$

The process $\{B_s - (s/t)B_t, s \in [0, t]\}$ is Brownian bridge with duration t seconds, i.e., Brownian motion starting from 0 and conditioned to be at 0 at time t . If \tilde{T}_r denotes the

hitting time of r by the absolute value of the Brownian bridge then we have a formula analogous to (4.18), $\limsup_{t/r^2 \rightarrow 0} (2t/r^2) \log \mathbf{P}(\tilde{T}_r < t) \leq -1$. Thus

$$\limsup_{t \rightarrow 0} t \log \mathbf{P}(A_2^c(t)) \leq -\varepsilon_9^2/2. \quad (4.20)$$

Note that Brownian motion $B = \{B_t, t \geq 0\}$ is a Gaussian process. Since for every $0 \leq s < t$, $B_s - (s/t)B_t$ and B_t have zero covariance, the process $\{B_s - (s/t)B_t, s \in [0, t]\}$ is independent of B_t . Similar remarks apply to W . Hence

$$\mathbf{P}(A_1(t) \cap A_2(t)) = \mathbf{P}(A_1(t))\mathbf{P}(A_2(t)) \quad \text{and} \quad \mathbf{P}(A_3(t) \cap A_4(t)) = \mathbf{P}(A_3(t))\mathbf{P}(A_4(t)).$$

This, (4.19) and (4.20) imply that

$$\mathbf{P}(A_5(t_1)) > p_1 \quad \text{for some } t_1, p_1 > 0. \quad (4.21)$$

Step 3. Let $R_t = \int_0^t \mathbf{n}(X_s) dL_s^X$. We will show that if $A_1(t_0) \cap A_2(t_0)$ holds for some $t_0 > 0$ then $|R_s| \leq 4\varepsilon_9$ for every $s \in [0, t_0]$.

Since $A_1(t_0)$ and $A_2(t_0)$ hold, $|B_t| \leq \varepsilon_3/4 + 2\varepsilon_9 \leq \varepsilon_3/2$ for every $t \leq t_0$. Thus when $\mathbf{d}(X_0, \partial D) \geq \varepsilon_3$, $X_0 + B_s \in \overline{D}$ for all $s \in [0, t_0]$. By the uniqueness of the solution to (4.1), $X_s = X_0 + B_s$ for $s \in [0, t_0]$, and $R_s = 0$ for all $s \in [0, t_0]$.

Suppose that $\mathbf{d}(X_0, \partial D) \leq \varepsilon_3$. By the assumptions made in Step 1(a), there exists a unique point $y \in \partial D$ with the smallest distance to X_0 . Let CS_1 be the orthonormal coordinate system such that $y = 0 \in \partial D$ and $\mathbf{n}(0)$ lies on the second axis. Recall from Step 1(a) that $\mathbf{n}(x) = (\mathbf{n}_1(x), \mathbf{n}_2(x))$ and $|\mathbf{n}_1(x)| \leq \mathbf{n}_2(x)/100$ for $x \in \partial D \cap \mathcal{B}(0, 3\varepsilon_3)$ in CS_1 . Write $R_t = (R_t^1, R_t^2)$. By the opening remarks in Step 1(c), $R_t \in \mathcal{L}(0, r)$ if $X_s \in \mathcal{B}(0, r)$ for all $s \leq t$. This implies that $|R_t^1| \leq R_t^2/100$ if $X_s \in \mathcal{B}(0, 3\varepsilon_3)$ for all $s \leq t$.

Let $T_1 = \inf\{t \geq 0 : |R_t| > 4\varepsilon_9\}$. We will assume that $T_1 < t_0$ and show that this leads to a contradiction. Since $|B_t| \leq \varepsilon_3/2$ for every $t \leq t_0$, we have $|B_t + R_t| \leq \varepsilon_3/2 + 4\varepsilon_9 \leq \varepsilon_3$ for every $t \leq T_1$. Hence,

$$|X_t| \leq |X_0| + |B_t + R_t| \leq \varepsilon_3 + \varepsilon_3 = 2\varepsilon_3 \quad \text{for every } t \leq T_1.$$

It follows that $R_t^2 \geq 0$ for every $t \leq T_1$ and $|R_{T_1}^1| \leq R_{T_1}^2/100$. Note that, by Step 1(a), the slope of the tangent line at points in $\mathcal{B}(0, 3\varepsilon_3) \cap \partial D$ is between $-1/100$ and $1/100$ and that $\mathbf{d}(X_{T_1}, X_0 + B_{T_1}) = |R_{T_1}| = 4\varepsilon_9$. The last observation and the fact that $X_{T_1} \in \partial D$ imply that $\mathbf{d}(X_0 + B_{T_1}, \partial D) \geq (2/3)|R_{T_1}| = 8\varepsilon_9/3$. Since $A_1(t_0)$ and $A_2(t_0)$

hold, $\mathbf{d}(X_0 + B_t, D) \leq 2\varepsilon_9$ for every $t \leq t_0$ and in particular for $t = T_1$. This contradiction proves the claim that $|R_s| \leq 4\varepsilon_9$ for every $s \in [0, t_0]$.

Let $\tilde{R}_t = \int_0^t \mathbf{n}(Y_s) dL_s^Y$. We claim that if $A_3(t_0) \cap A_4(t_0)$ holds for some $t_0 > 0$ then $|\tilde{R}_s| \leq \varepsilon_3$ for every $s \in [0, t_0]$. To see this, observe that $\mathbf{d}(Y_0 + W_t, D) \leq \varepsilon_3/4 + 2\varepsilon_9 \leq \varepsilon_3/2$ for $t \leq t_0$ and use that same argument as in the case of R_t .

Step 4. Fix $t_0 > 0$. We will show that $A_5(t_0) \subset F_5(t_0)$. Assume that $A_5(t_0)$ holds.

Since $A_1(t_0)$ holds, we have in view of Step 3,

$$\begin{aligned} & \mathbf{d}(X_{t_0}, X_0 + (\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0)) \\ & \leq \mathbf{d}(B_{t_0}, (\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0)) + |R_{t_0}| \leq \varepsilon_9 + 4\varepsilon_9 \leq \varepsilon_6 \wedge \varepsilon_8. \end{aligned}$$

In other words, $F_3(t_0)$ holds.

Since $A_1(t_0)$ and $A_3(t_0)$ hold we have, using simple geometry,

$$\mathbf{d}(X_0 + B_{t_0}, Y_0 + W_{t_0}) \leq \mathbf{d}(X_0, Y_0) + 3\varepsilon_9. \quad (4.22)$$

We will apply (4.10) with $x_1 = X_0 + B_{t_0}$, $y_1 = Y_0 + W_{t_0}$, $\mathbf{v} = R_{t_0}$ and $\mathbf{w} = \tilde{R}_{t_0}$. We have assumed that $A_5(t_0)$ holds so $|B_t| \leq \varepsilon_3$ and $|W_t| \leq \varepsilon_3$ for $t \leq t_0$. This and Step 3 imply that for $t \leq t_0$,

$$\mathbf{d}(X_t, X_0) \leq |B_t| + |R_t| \leq 2\varepsilon_3,$$

and similarly $\mathbf{d}(Y_t, Y_0) \leq 2\varepsilon_3$. By the opening remarks in Step 1(c),

$$\mathbf{v} = \int_0^{t_0} \mathbf{n}(X_t) dL_t^X \in \mathcal{L}(X_0, 2\varepsilon_3) \quad \text{and} \quad \mathbf{w} = \int_0^{t_0} \mathbf{n}(Y_t) dL_t^Y \in \mathcal{L}(Y_0, 2\varepsilon_3).$$

By Step 3, $|\mathbf{v}| \leq \varepsilon_3$ and $|\mathbf{w}| \leq \varepsilon_3$. We have shown that all the conditions listed in Step 1(c) are satisfied so we can apply (4.10) to obtain

$$\mathbf{d}(X_{t_0}, Y_{t_0}) \leq \mathbf{d}(X_0, Y_0) + 3\varepsilon_9.$$

This proves that $F_4(t_0)$ holds.

It will now suffice to show that if $F_2(t_0)$ does not hold then $F_1(t_0)$ does. Assume that $F_2(t_0)$ does not hold.

If all of the following conditions hold,

$$|B_{t_0} - (\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0)| \leq \varepsilon_6/2 \wedge c_8, \quad (4.23)$$

$$|W_{t_0} - (\varepsilon_3/4)(Y_0 - X_0)/\mathbf{d}(X_0, Y_0)| \leq \varepsilon_6/2, \quad (4.24)$$

$$\mathbf{d}(X_{t_0}, X_0 + B_{t_0}) \leq \varepsilon_6/2, \quad (4.25)$$

$$\mathbf{d}(Y_{t_0}, Y_0 + W_{t_0}) \leq \varepsilon_6/2, \quad (4.26)$$

then (4.12) and (4.13) hold with $(x_1, y_1) = (X_{t_0}, Y_{t_0})$ and $(x_0, y_0) = (X_0, Y_0)$, and this implies $F_2(t_0)$, which is a contradiction. Hence, at least one of the conditions (4.23)-(4.26) must fail. The first of these conditions holds because $A_1(t_0)$ is true. By Step 3, (4.25) holds.

Suppose that (4.26) fails. In view of (4.22), we can apply (4.17) to $x_1 = X_0 + B_{t_0}$, $y_1 = Y_0 + W_{t_0}$, $\mathbf{v} = R_{t_0}$ and $\mathbf{w} = \tilde{R}_{t_0}$ to obtain

$$\mathbf{d}(X_{t_0}, Y_{t_0}) \leq \mathbf{d}(X_0, Y_0) + 3\varepsilon_9 - 2\varepsilon_7 \leq \mathbf{d}(X_0, Y_0) - \varepsilon_7.$$

Hence, we have $F_1(t_0)$ in this case.

Suppose that (4.24) fails. Then, in view of Step 3, (4.14)-(4.16) hold with $(x_0, y_0) = (X_0, Y_0)$, $(x_1, y_1) = (X_0 + B_{t_0}, Y_0 + W_{t_0})$, $\mathbf{v} = R_{t_0}$ and $\mathbf{w} = \tilde{R}_{t_0}$ and we have

$$\mathbf{d}(X_{t_0}, Y_{t_0}) \leq \mathbf{d}(X_0, Y_0) - \varepsilon_7.$$

Hence, $F_1(t_0)$ holds. This proves that $A_5(t_0) \subset F_5(t_0)$.

Step 5. Fix some $\varepsilon_1 \in (0, \varepsilon_0)$ and let ε_j 's be defined relative to ε_1 as in Step 1. Let ρ be the diameter of D and let N_0 be an integer greater than $4\rho/\varepsilon_7$. Recall from (4.21) in Step 2 that for some $p_1, t_1 > 0$ we have $\mathbf{P}(A_5(t_1)) > p_1$. Let

$$S_1 = \inf\{t \geq 0 : \mathbf{d}(X_t, Y_t) \leq \varepsilon_1 \text{ or } F_5(t) \text{ holds}\} \wedge (2t_1).$$

By (4.21) and Step 4, $\mathbf{P}(S_1 \leq t_1) > p_1$. Recall that θ stands for the usual Markov shift and define

$$S_0 = 0 \quad \text{and} \quad S_k = S_1 \circ \theta_{S_{k-1}} + S_{k-1} \quad \text{for } k \geq 1.$$

Recall integer N defined in Step 1. By the strong Markov property, with probability no less than $p_1^{2NN_0} > 0$, we have $S_k - S_{k-1} \leq t_1$ for all $k \leq 2NN_0$.

We will argue that if $\bigcap_{k \leq 2NN_0} \{S_k - S_{k-1} \leq t_1\}$ holds then $\mathbf{d}(X_t, Y_t) \leq \varepsilon_1$ for some $t \leq 2NN_0 t_1$. Assume otherwise. Then $F_5(S_1) \circ \theta_{S_{k-1}}$ holds for every $k \leq 2NN_0$. In particular, $F_4(S_1) \circ \theta_{S_{k-1}}$ holds for every $k \leq 2NN_0$. Let $F_6(t) = F_2(t) \cap F_3(t)$. Since $F_5(t) \subset F_4(t) \cap (F_1(t) \cup (F_2(t) \cap F_3(t)))$, for every $k \leq 2NN_0$, at least one of the events $F_1(S_1) \circ \theta_{S_{k-1}}$ and $F_6(S_1) \circ \theta_{S_{k-1}}$ holds.

Consider any $j \leq N_0$. If $F_6(S_1) \circ \theta_{S_{k-1}}$ holds for $k = 2jN, 2jN + 1, \dots, 2(j+1)N - 1$ then X_{S_k} 's and Y_{S_k} 's satisfy the following conditions for $k = 2jN, 2jN + 1, \dots, 2(j+1)N - 1$,

$$\begin{aligned} \mathbf{d}(X_{S_{k+1}}, X_{S_k} + (\varepsilon_3/4)(Y_{S_k} - X_{S_k})/\mathbf{d}(X_{S_k}, Y_{S_k})) &\leq \varepsilon_6, \\ |\angle(X_{S_k} - Y_{S_k}, X_{S_{k+1}} - Y_{S_{k+1}})| &\leq \varepsilon_5. \end{aligned}$$

This implies, by Step 1(e), that for $k = 2jN, 2jN + 1, \dots, 2(j+1)N - 2$,

$$|\angle(X_{S_{k+1}} - X_{S_k}, X_{S_{k+2}} - X_{S_{k+1}})| \leq \varepsilon_4.$$

Since $F_3(S_1) \circ \theta_{S_{k-1}}$ holds, we also have $\mathbf{d}(X_{S_{k+1}}, X_{S_k}) \geq \varepsilon_3/8$ for the same range of k (to see this, recall from Step 1(e) that $\varepsilon_6 < \varepsilon_3/8$). Hence $X_{S_{2(j+1)N-2}}$ must be outside \overline{D} , according to the definition of ε_4 and N in Step 1(d). Since X always stays inside \overline{D} , at least one of the events $F_6(S_1) \circ \theta_{S_{k-1}}$ must fail for some $2jN \leq k \leq 2(j+1)N - 1$. Hence, at least one event $F_1(S_1) \circ \theta_{S_{k-1}}$ holds for some $2jN \leq k \leq 2(j+1)N - 1$. Since $F_4(S_1) \circ \theta_{S_{k-1}}$ holds for every $k \leq 2NN_0$, there is a reduction of at least $\varepsilon_7/2$ in the distance between X and Y on every interval $[S_{2jN}, S_{2(j+1)N}]$, that is,

$$\mathbf{d}(X_{S_{2jN}}, Y_{S_{2jN}}) \leq \mathbf{d}(X_{S_{2(j-1)N}}, Y_{S_{2(j-1)N}}) - \frac{\varepsilon_7}{2} \quad \text{for every } j \in \{1, \dots, N_0\}.$$

Summing over j we obtain

$$\mathbf{d}(X_{S_{2N_0N}}, Y_{S_{2N_0N}}) \leq \mathbf{d}(X_0, Y_0) - \frac{N_0\varepsilon_7}{2} \leq \mathbf{d}(X_0, Y_0) - 2\rho < 0.$$

This contradiction proves our claim that

$$\text{if } \bigcap_{k \leq 2NN_0} \{S_k - S_{k-1} \leq t_1\} \text{ holds, then } \mathbf{d}(X_t, Y_t) \leq \varepsilon_1 \text{ for some } t \leq 2t_1NN_0.$$

We have shown that $\mathbf{d}(X_t, Y_t) \leq \varepsilon_1$ for some $t \leq 2t_1NN_0$ with probability greater than $p_2 \stackrel{\text{df}}{=} p_1^{2NN_0} > 0$. By the Markov property, $\mathbf{d}(X_t, Y_t) \leq \varepsilon_1$ for some $t \leq 2jt_1NN_0$ with probability greater than $1 - (1 - p_2)^j$. To complete the proof, it suffices to let $j \rightarrow \infty$. \square

Two of the assumptions on the boundary of D made in Theorem 4.3, that it is convex with C^1 -smooth boundary and it does not contain any line segments, are convenient from the technical point of view but most likely one can dispose of them with analysis more refined than that in our proof.

Example 4.4. Suppose that D is the annulus $\{x \in \mathbb{R}^2 : 1 < |x| < 2\}$. The rotation of D around $(0, 0)$ with an angle in $(0, 2\pi)$ is an isometry with no fixed points. Hence, there exists a shy coupling of reflected Brownian motions in this annulus (see Example 3.3).

There are many open problems concerning existence of shy couplings but we find the following two questions especially intriguing. Recall that $\mathcal{B}(x, r)$ denotes the open ball with center x and radius r .

Open problems 4.5. (i) Does there exist a shy coupling of reflected Brownian motions in $\mathcal{B}((0,0),3) \setminus \mathcal{B}((1,0),1)$?

(ii) Does there exist a shy coupling of reflected Brownian motions in any simply connected planar domain?

We end this paper with a vague remark concerning a potential relationship between shy couplings and an old and well known problem of “fixed points.” Suppose that \mathcal{S} is a topological space. If every continuous mapping $I : \mathcal{S} \rightarrow \mathcal{S}$ has a fixed point, i.e., a point $x \in \mathcal{S}$ such that $I(x) = x$, then we say that \mathcal{S} has the fixed point property. One of the most famous fixed point theorems is that of Brouwer—it asserts that a closed ball in \mathbb{R}^d has the fixed point property. Spheres obviously do not have the fixed point property. Some of our results may suggest that a shy coupling exists if and only if the state space does not have the fixed point property. Example 3.8 applied to the graph illustrated in Fig. 3.6 shows that this conjecture is false at this level generality. It is possible, though, that a weaker form of this assertion is true—we leave it as an open problem.

REFERENCES

- [BC1] K. Burdzy and Z.-Q. Chen, Local time flow related to skew Brownian motion. *Ann. Probab.* **29** (2001), 1693-1715.
- [BC2] K. Burdzy and Z.-Q. Chen, Coalescence of synchronous couplings. *Probab. Theory Rel. Fields* **123** (2002), 553–578.
- [BCJ] K. Burdzy, Z.-Q. Chen and P. Jones, Synchronous couplings of reflected Brownian motions in smooth domains. Preprint, 2005.
- [BK] K. Burdzy and H. Kaspi, Lenses in skew Brownian flow. *Ann. Probab.* **32** (2004), 3085–3115.
- [FW] M. Freidlin and A. Wentzell, Diffusion processes on graphs and the averaging principle. *Ann. Probab.* **21** (1993), 2215-2245.
- [F] A. Friedman, *Differential Games*. Wiley, New York, 1971.
- [HS] J.M. Harrison and L.A. Shepp, On skew Brownian motion. *Ann. Probab.* **9** (1981), 309-313.

- [KS] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Second edition. Springer, New York, 1991.
- [L] T. Lindvall, *Lectures on the Coupling Method*. Wiley, New York, 1992.
- [LS] P. L. Lions and A. S. Sznitman, Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37** (1984), 511-537.
- [RY] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*. Springer, New York, 1991.
- [Ta] H. Tanaka, Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.* **9** (1979), 163-177.

I.B.: Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
itai.benjamini@weizmann.ac.il

K.B. and Z.C.: Department of Mathematics, Box 354350, University of Washington, Seattle, WA 98115-4350, USA
burdzy@math.washington.edu, zchen@math.washington.edu