

**Original citation:**

Banerjee, S. & Kendall, W.S. Probabability Theory Related Fields (2016). doi:10.1007/s00440-016-0706-4

**Permanent WRAP URL:**

<http://wrap.warwick.ac.uk/69885>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: <http://creativecommons.org/licenses/by/4.0/>

**A note on versions:**

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk)

# Rigidity for Markovian maximal couplings of elliptic diffusions

Sayan Banerjee<sup>1</sup> · Wilfrid S. Kendall<sup>1</sup>

Received: 8 December 2014 / Revised: 26 March 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Maximal couplings are (probabilistic) couplings of Markov processes such that the tail probabilities of the coupling time attain the total variation lower bound (Aldous bound) uniformly for all time. Markovian (or immersion) couplings are couplings defined by strategies where neither process is allowed to look into the future of the other before making the next transition. Markovian couplings are typically easier to construct and analyze than general couplings, and play an important role in many branches of probability and analysis. Hsu and Sturm, in a preprint circulating in 2007, but later published in 2013, proved that the *reflection-coupling* of Brownian motion is the *unique Markovian maximal coupling (MMC)* of Brownian motions starting from two different points. Later, Kuwada (Electron J Probab 14(25), 633–662, 2009) proved that the existence of a MMC for Brownian motions on a Riemannian manifold enforces existence of a reflection structure on the manifold. In this work, we investigate suitably regular elliptic diffusions on manifolds, and show how consideration of the diffusion geometry (including dimension of the isometry group and flows of isometries) is fundamental in classification of the space and the generator of the diffusion for which an MMC exists, especially when the MMC also holds under local perturbations of the starting points for the coupled diffusions. We also describe such diffusions in terms of *Killing vectorfields* (generators of isometry groups) and *dilation vectorfields* (generators of scaling symmetry groups). This permits a complete characterization of those possible manifolds and their diffusions for which there exists

---

This work was supported by EPSRC Research Grant EP/K013939.

---

✉ Sayan Banerjee  
sayan.banerjee20@gmail.com

Wilfrid S. Kendall  
W.S.Kendall@warwick.ac.uk

<sup>1</sup> University of Warwick, Coventry, England

a MMC under local perturbations of the starting points of the coupled diffusions. For example, in the time-homogeneous case it is shown that the only possible manifolds that may arise are Euclidean space, hyperbolic space and the hypersphere. Moreover the permissible drifts can then derive only from rotation isometries of these spaces (and dilations, in the Euclidean case). In this sense, a geometric rigidity phenomenon holds good.

**Keywords** Characteristic operator · Coupling · Diffusion-geodesic completeness · Diffusion geometry · Diffusion matrix · Elliptic diffusion · Global isometry group · Homogeneous space · Immersion coupling · Infinitesimal generator · Involutive isometry · Killing vectorfield · Laplace–Beltrami operator · Large deviations · Local perturbation condition · Markovian coupling · Maximal coupling · Maximally symmetric space · Orthonormal frame bundle · Riemannian manifold · Reflection coupling · Rigidity · Stochastic completeness · Stochastic differential equation · Stochastic parallel transport · Stratonovich differential · Strong maximum principle · Topogonov comparison theorem · Totally geodesic submanifold

**Mathematics Subject Classification** 60G05 · 58J65 · 60J60

## 1 Introduction

Let  $(\Omega_1, \mathcal{F}_1, \mu_1)$  and  $(\Omega_2, \mathcal{F}_2, \mu_2)$  be two probability spaces. A (probabilistic) *coupling* of  $\mu_1$  and  $\mu_2$  is a measure  $\mu$  on the product measurable space  $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$  with marginals  $\mu_1$  and  $\mu_2$ . This paper considers the question of coupling of (the laws of) two realizations  $X$  and  $Y$  of a Markov process on some state space  $S$ . We distinguish two important classes. The first class (thematic for the foundational theory of probabilistic coupling) consists of couplings where, with positive probability,  $X$  and  $Y$  can stick together and move as a single process after some random time

$$\tau = \inf\{s > 0 : X_t = Y_t \text{ for all } t > s\};$$

here  $\tau$  is called the *coupling time*. The other class consists of couplings (*Shy Couplings*) where the two processes  $X$  and  $Y$  remain separated by at least a fixed positive distance  $\varepsilon$  for all time. Recent investigations of the second class of couplings can be found in [1] and [2, 3]; in this article, we concentrate on the first class.

Probabilistic coupling is a central technique of modern probability theory [27, 38]. Attention naturally focusses on a fundamental question: how fast can we make  $X$  and  $Y$  meet? This has direct relevance, for example to the study of probabilistic algorithms and to gradient estimates for harmonic functions, and is also very valuable in eliciting the range of possibilities for coupling constructions. Mathematically, this amounts to constructing couplings where  $\mathbb{P}[\tau > t]$  is minimised for all time  $t$ . The Aldous inequality states that, for any  $t > 0$ ,

$$\mathbb{P}[\tau > t] \geq \|\mu_{1,t} - \mu_{2,t}\|_{TV}, \quad (1)$$

where  $\mu_{1,t}$  and  $\mu_{2,t}$  are the distributions of  $X_t$  and  $Y_t$  respectively, while

$$\|v\|_{TV} = \sup\{|v(A)| : \text{measurable } A\}$$

denotes the total variation norm on signed measures  $v$ . Thus a maximally efficient possible coupling (a *Maximal Coupling*) would attain equality in the Aldous inequality (1) for all times  $t > 0$ , thus solving a multi-objective optimization problem. The remarkable construction of [16], later simplified in a most elegant way by [33], shows that maximal couplings always exist for discrete Markov chains. [15] generalized the construction to the case of non-Markovian processes; [37] generalized it to continuous-time càdlàg processes. Here is a summary of the Pitman approach, which is a model for the construction below (in Sect. 1.1) of maximal couplings of diffusions. A deterministic time-varying interface is constructed using the transition probabilities of the diffusions which are to be coupled. The distribution of the coupling time is elicited using the deficits of the transition probability masses integrated on each side of the interface (at any particular time, these deficits are equal and correspond to the probability of one, equivalently both, of the coupled processes hitting the interface at this time). Now, the coupling time is sampled from this distribution, and the coupling location corresponds to a point on the interface at this time. Finally, the coupling is realized by constructing a single process forward in time and time-reversed time-inhomogeneous diffusions connecting starting locations to the location and moment of coupling, conditioning to avoid hitting the interface prematurely.

The major drawback of all these constructions is they are typically very implicit; in most cases, it is extremely hard, if not impossible, to make detailed calculations for such couplings. This is a strong motivation for considering Markovian couplings, which we now describe.

Let  $X$  and  $Y$  be Markov processes starting from  $x_0$  and  $y_0$  respectively. Let  $\mathcal{F}_s = \sigma\{(X_{s'}, Y_{s'}) : s' \leq s\}$  denote the joint filtration generated by  $X$  and  $Y$  together up to time  $s$ . A coupling of  $X$  and  $Y$  is called *Markovian* if the joint process

$$\{(X_{t+s}, Y_{t+s}) : t \geq 0\} \text{ conditioned on } \mathcal{F}_s$$

is again a coupling of the laws of  $X$  and  $Y$ , but now starting from  $(X_s, Y_s)$ . (An alternative martingale-based characterization makes a succinct connection to the theory of immersions of filtrations. For this reason Markovian couplings are also called *immersion couplings*: [22])

A natural and immediate question is, when can a maximal coupling of two diffusions be Markovian? The standard (and elegant) example in the literature is the *reflection-coupling* of Euclidean Brownian motions starting from two different points: the second Brownian path is obtained from the first by reflecting the first path on the hyperplane bisecting the line joining the starting points until the first path (equivalently, the second, reflected, path) hits this hyperplane. Both paths then evolve together (“synchronously”) as a single Brownian path. Straightforward calculations, based on the reflection principle, show that this construction is in fact a Markovian maximal coupling (MMC). Furthermore, [18] proved that this is the *unique* such coupling for Euclidean Brownian motion. A few other examples are discussed in the literature:

Ornstein Uhlenbeck processes [9], also Brownian motion on manifolds which possess certain reflection symmetries. The reflection coupling idea manifests itself throughout the area of probabilistic coupling: for example it has a natural generalization to Brownian motion on Riemannian manifolds [10, 19], involving stochastic parallel transport and development, and not requiring any symmetries of the manifold. However it seems unlikely that such generalizations will normally provide maximal couplings. Kuwada [25] investigated this question for Brownian motion on manifolds (and their generalisations to metric spaces). Under suitable mild regularity assumptions he showed that a reflection symmetry of the space is *necessary* for the existence of a Markovian maximal coupling of two Brownian motions started from a specified pair of points. Working under some further assumptions, he proved that the fixed point set of the symmetry (the “mirror”, characterizing this isometry) does not change with time; the maximal coupling is given simply by reflecting one process onto the other using the reflection symmetry defined by this mirror.

The aim of this paper is to develop the results of Kuwada to the case of general regular elliptic diffusions with smooth coefficients. It will be shown that Markovian maximal couplings are rare, in the sense that a stable local existence result enforces extreme global symmetry on the manifold: a kind of rigidity result. Section 2 considers implications of existence of Markovian maximal couplings for  $d$ -dimensional Euclidean diffusions (“Euclidean” here meaning that the diffusion matrix is the identity matrix), under rather general regularity assumptions on the (possibly time-inhomogeneous) drift. Extending Kuwada’s argument, the existence of an MMC implies there is a mirror symmetry between the coupled processes at any given time. However the influence of the non-zero drift now means that the mirror can vary deterministically with time, making the coupled dynamics considerably more complicated. We study the evolution of the mirror in time using stochastic calculus and we obtain a functional equation that the drift must satisfy for a Markovian maximal coupling to exist. This equation can be used to characterise all time-inhomogeneous diffusions which admit such couplings.

In the time-homogeneous case the characterization can be refined under the additional hypothesis that there is also a Markovian maximal coupling under local perturbation of the starting points, which is to say, Markovian maximal couplings exist locally in a stable sense:

**Definition 1** (*Local Perturbation Condition (LPC)*) There is  $r > 0$ , and initial points  $\mathbf{x}_0$  and  $\mathbf{y}_0$ , such that there exists a Markovian maximal coupling of the diffusion processes  $X$  and  $Y$  starting from  $\mathbf{x}$  and  $\mathbf{y}$  for every  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$ , where  $\mathcal{B}(\mathbf{x}_0, r)$  is the open metric ball centred at  $\mathbf{x}_0$  and of radius  $r$ .

We will show that, for any dimension  $d \geq 1$ , *LPC* holds for a suitably regular Euclidean diffusion with time-homogeneous drift if and only if the drift takes the form  $\mathbf{b}(\mathbf{x}) = \lambda \mathbf{x} + T\mathbf{x} + \mathbf{c}$ , where  $\lambda$  is a scalar,  $T$  is a skew-symmetric matrix and  $\mathbf{c}$  is a fixed vector. This implies that *Brownian motion with constant drift* and *Ornstein–Uhlenbeck process* are the only *one-dimensional* examples of time-homogeneous diffusions for which there are successful Markovian maximal couplings from arbitrary pairs of starting points. In higher dimensions, for regular Euclidean diffusions under *LPC*, essentially the same

is true except that the drift may also include a rotational component. In one dimension, even without *LPC*, it turns out that a Markovian maximal coupling exists between two copies of a regular diffusion started from  $x_0$  and  $y_0$  if and only if the drift is either affine or an odd function around the midpoint of the starting points.

Section 3 considers Markovian maximal couplings of Brownian motion with time-homogeneous drift on a complete Riemannian manifold  $M$  under *LPC*. This is the natural generalization of the context of Sect. 2, since a regular elliptic diffusion on Euclidean space furnishes the space with a Riemannian metric by means of inverting the diffusion matrix, and then the diffusion is converted into a Brownian motion with drift on the resulting Riemannian manifold, so that the Riemannian geometry serves to classify a variety of diffusions (compare the rather similar role of Fisher information in theoretical statistics). We assume that the elliptic diffusion is stochastically complete, and also diffusion-geodesically complete, in the sense that the diffusion Riemannian geometry is geodesically complete. Strikingly, *LPC* then produces a geometric rigidity phenomenon, namely a complete classification of the space  $M$  as one of the three model spaces  $\mathbb{R}^d$  (Euclidean space),  $\mathbb{S}^d$  (Sphere) and  $\mathbb{H}^d$  (Hyperbolic space) depending upon the sign of the (necessarily constant) curvature  $K$  (see Theorem 38 in Sect. 3). The Euclidean case is fully covered in Sect. 2, and delivers the necessary ideas and techniques which we generalise to the manifold setup in Sect. 3 to study Markovian maximal couplings on the other two spaces. It turns out that the only drifts which can yield Markovian maximal couplings are given by the *Killing vectorfields*, defined as infinitesimal generators for the rigid motion group (namely, generators of one-parameter subgroups of isometries).

In this paper we confine our considerations to the case of elliptic diffusions, where there is a strong connection to Riemannian geometry, and path-continuity permits the formation of interfaces of co-dimension 1 separating pairs of initial points. Possible extensions to hypoelliptic diffusions or to general Markov chains are potentially of great interest, but we leave these questions as topics for future work.

## 1.1 Markovian maximal couplings: general properties

We complete this introduction by defining some general notation and by describing some basic general properties of Markovian maximal couplings for general Markov processes on a metric space  $(M, \text{dist})$ . Kuwada [25] derived results similar to Lemmas 2 and 3 below. For the sake of clearer exposition, and as we are primarily interested in diffusion processes, we will state the results for continuous-time Markov processes. Denote the Markov process under consideration by  $X$ .

We assume that the metric space supports a positive Borel measure  $m$  with  $0 < m(B) < \infty$  for any metric ball  $B$  of finite radius. Consequently, the closed support of  $m$  is the whole of  $M$ . We further assume that for any  $t > s \geq 0$ , the conditional distribution law  $\mathcal{L}(X_t | X_s = x)$  is absolutely continuous with respect to  $m$  and has a probability kernel density given by  $p(s, \mathbf{x}; t, \mathbf{z})$  for  $\mathbf{x}, \mathbf{z} \in M$  and  $0 \leq s < t$ .

Let  $\mu$  denote the law of a Markovian maximal coupling  $(X, Y)$  of two copies of our Markov process started from  $(\mathbf{x}_0, \mathbf{y}_0)$ , which can be thought of as a measure on the coupled path-space  $C[0, \infty)^2$ , and let

$$\tau = \inf\{s > 0 : X_t = Y_t \text{ for all } t > s\}$$

denote the coupling time of  $X$  and  $Y$ .

Motivated by Pitman's construction for finite Markov chains, we write

$$\alpha(s, \mathbf{x}, \mathbf{y}, t, \mathbf{z}) = p(s, \mathbf{x}; t, \mathbf{z}) - p(s, \mathbf{y}; t, \mathbf{z}),$$

and set  $\alpha^+(s, \mathbf{x}, \mathbf{y}, t, \mathbf{z}) = \max(\alpha(s, \mathbf{x}, \mathbf{y}, t, \mathbf{z}), 0)$  and  $\alpha^-(s, \mathbf{x}, \mathbf{y}, t, \mathbf{z}) = \max(-\alpha(s, \mathbf{x}, \mathbf{y}, t, \mathbf{z}), 0)$ . If  $s = 0$  (and thus  $\mathbf{x} = \mathbf{x}_0$  and  $\mathbf{y} = \mathbf{y}_0$ ), then we abbreviate  $\alpha(t, \mathbf{z})$  for  $\alpha(s, \mathbf{x}_0, \mathbf{y}_0, t, \mathbf{z})$  and similarly for other quantities.

We will be dealing with Markov processes which are possibly time-inhomogeneous, so we say a Markov process starts from  $(t, \mathbf{x})$  if we are looking at the distribution law  $\mathcal{L}(\theta_t X \mid X_t = \mathbf{x})$ , where  $\theta$  denotes the time-shift operator given by  $(\theta_t X)_s = X_{t+s}$ .

Define the *interface* between  $p(0, \mathbf{x}_0; \cdot, \cdot)$  and  $p(0, \mathbf{y}_0; \cdot, \cdot)$  at time  $t$  to be the region where the corresponding heat kernels agree:

$$I(\mathbf{x}_0, \mathbf{y}_0, t) = \{\mathbf{z} \in M : p(0, \mathbf{x}_0; t, \mathbf{z}) = p(0, \mathbf{y}_0; t, \mathbf{z})\}. \quad (2)$$

Also write

$$\begin{aligned} I^-(\mathbf{x}_0, \mathbf{y}_0, t) &= \{\mathbf{z} \in M : p(0, \mathbf{x}_0; t, \mathbf{z}) > p(0, \mathbf{y}_0; t, \mathbf{z})\}, \\ I^+(\mathbf{x}_0, \mathbf{y}_0, t) &= \{\mathbf{z} \in M : p(0, \mathbf{x}_0; t, \mathbf{z}) < p(0, \mathbf{y}_0; t, \mathbf{z})\}. \end{aligned} \quad (3)$$

Finally, define the perpendicularly bisecting set (or “hyperplane”) and the associated “half-spaces” (note that these are indeed a hyperplane and half-spaces in the Euclidean case):

$$\begin{aligned} H(\mathbf{x}, \mathbf{y}) &= \{\mathbf{z} \in M : \text{dist}(\mathbf{x}, \mathbf{z}) = \text{dist}(\mathbf{y}, \mathbf{z})\}, \\ H^-(\mathbf{x}, \mathbf{y}) &= \{\mathbf{z} \in M : \text{dist}(\mathbf{x}, \mathbf{z}) < \text{dist}(\mathbf{y}, \mathbf{z})\}, \\ H^+(\mathbf{x}, \mathbf{y}) &= \{\mathbf{z} \in M : \text{dist}(\mathbf{x}, \mathbf{z}) > \text{dist}(\mathbf{y}, \mathbf{z})\}. \end{aligned} \quad (4)$$

**Lemma 2** *Any joint maximal coupling law can be related to differences of the transition probability kernel densities as follows: for any Borel subset  $A$  of  $M$ , and  $s > 0$ ,*

$$\begin{aligned} \mu(X_s \in A, \tau > s) &= \int_A \alpha^+(s, \mathbf{x}) m(d\mathbf{x}), \\ \mu(Y_s \in A, \tau > s) &= \int_A \alpha^-(s, \mathbf{x}) m(d\mathbf{x}). \end{aligned}$$

*Proof* It is immediate that  $\mu(X_s \in A, \tau \leq s) \leq \mu(X_s \in A)$ . If  $p(0, \mathbf{x}_0; s, \cdot) \leq p(0, \mathbf{y}_0; s, \cdot)$  on  $A$  then

$$\begin{aligned}
 \mu(X_s = Y_s \in A, \tau \leq s) &= \mu(X_s \in A, \tau \leq s) \leq \mu(X_s \in A) \\
 &= \int_A p(0, \mathbf{x}_0; s, \mathbf{x}) m(d\mathbf{x}) \\
 &= \int_A p(0, \mathbf{x}_0; s, \mathbf{x}) \wedge p(0, \mathbf{y}_0; s, \mathbf{x}) m(d\mathbf{x}).
 \end{aligned}$$

Interchanging the rôles of  $X$  and  $Y$ , a corresponding argument applies if  $p(0, \mathbf{x}_0; s, \cdot) \geq p(0, \mathbf{y}_0; s, \cdot)$  on  $A$ . Hence additivity shows that for all  $A$  the coupling must satisfy

$$\mu(X_s = Y_s \in A, \tau \leq s) \leq \int_A p(0, \mathbf{x}_0; s, \mathbf{x}) \wedge p(0, \mathbf{y}_0; s, \mathbf{x}) m(d\mathbf{x}). \quad (5)$$

Finally, Aldous' inequality (1) is by definition an equality for a maximal coupling, so

$$\mu(\tau \leq s) = \int_{\mathbb{R}^d} p(0, \mathbf{x}_0; s, \mathbf{x}) \wedge p(0, \mathbf{y}_0; s, \mathbf{x}) m(d\mathbf{x}). \quad (6)$$

It follows that the inequality (5) must in fact be an equality. This proves the lemma.  $\square$

Only maximality was required for Lemma 2. If in addition  $\mu$  is Markovian, then the conditional law  $\mathcal{L}(\theta_s X, \theta_s Y \mid \mathcal{F}_s)$  describes a Markovian coupling of two copies of our Markov process starting from  $((s, X_s), (s, Y_s))$ . Such a coupling therefore satisfies the following *flow property*:

**Lemma 3** *If  $\mu$  is a Markovian maximal coupling and  $\mu_s = \mathcal{L}(X_s, Y_s)$  then, for  $\mu_s$ -almost every  $(\mathbf{x}, \mathbf{y})$  with  $\mathbf{x} \neq \mathbf{y}$  the conditional law  $\mathcal{L}(\theta_s X, \theta_s Y \mid X_s = \mathbf{x}, Y_s = \mathbf{y})$  gives a Markovian maximal coupling of  $(X, Y)$  starting from  $((s, \mathbf{x}), (s, \mathbf{y}))$ .*

*Proof* This follows immediately from the maximality of  $\mu$  and the fact that  $\mu$  is Markovian.  $\square$

We now introduce notation to describe the set of pairs of initial points in the closed support of  $\mu_s$  for which the forward processes  $(\theta_s X, \theta_s Y)$  do indeed generate a maximal coupling:

$$\begin{aligned}
 \mathcal{M}(\mu_s) &= \{(\mathbf{x}, \mathbf{y}) \in \text{Support}(\mu_s) : \mathbf{x} \neq \mathbf{y} \text{ and } \mathcal{L}(\theta_s X, \theta_s Y \mid X_s = \mathbf{x}, Y_s = \mathbf{y}) \\
 &\quad \text{yields a maximal coupling of } (X, Y) \text{ starting from } ((s, \mathbf{x}), (s, \mathbf{y}))\}.
 \end{aligned}$$

We conclude this introduction by noting an elementary observation about couplings of Markov processes.

**Lemma 4** *For each  $t \geq 0$ , let  $F_t : (\Omega_1, \mathcal{F}_1) \rightarrow (\Omega_2, \mathcal{F}_2)$  be a bijective mapping between two measurable spaces such that  $F_t, F_t^{-1}$  are measurable. Then, for any Markov process  $\{X_t : t \geq 0\}$  on  $\Omega_1$ ,  $\{F_t(X_t) : t \geq 0\}$  defines a Markov process on  $\Omega_2$ . Furthermore  $\{(X_t, Y_t) : t \geq 0\}$  is a (Markovian) maximal coupling of Markov processes on  $\Omega_1$  if and only if  $\{(F_t(X_t), F_t(Y_t)) : t \geq 0\}$  is a (Markovian) maximal coupling on  $\Omega_2$ .*



*Proof* The first assertion is a direct consequence of the general definition of conditional expectation. The second assertion follows from the definition of maximality.  $\square$

## 2 Markovian maximal couplings on Euclidean spaces

We consider diffusions on Euclidean space  $\mathbb{R}^d$  with infinitesimal generator

$$L = \frac{1}{2} \sum_{i=1}^d \partial_i^2 + \sum_{i=1}^d b_i(t, \mathbf{x}) \partial_i, \quad (7)$$

where  $\partial_i = \frac{\partial}{\partial x_i}$ . In the following,  $X$  will be used to denote a diffusion with the above generator. We will refer below to such a diffusion as a *Euclidean diffusion*, because diffusions with general diffusion coefficients are covered in Sect. 3 as instances of ‘Brownian motion plus drift on a manifold’. We make the following very general regularity assumptions (not necessary for all of our results, but imposed globally to streamline the exposition):

- (A1) The drift vectorfield  $\mathbf{b} : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$  is continuously differentiable in the second (space) variable, moreover  $\mathbf{b}$  and all its first-order spatial partial derivatives  $\partial_i \mathbf{b}$  are bounded on compact subsets of  $[0, \infty) \times \mathbb{R}^d$ .
- (A2) For every  $t > s \geq 0$ , and  $\mathbf{x}, \mathbf{z} \in \mathbb{R}^d$ , the conditional distribution law  $\mathcal{L}(X_t | X_s = x)$  is the law of a diffusion with transition probability density kernel  $p(s, \mathbf{x}; t, \mathbf{z})$  (density with respect to Lebesgue measure), which is jointly continuous in all its arguments. Moreover,  $p(s, \cdot; \cdot, \cdot)$  is positive everywhere when  $s > 0$ . Finally, the density  $p(s, \mathbf{x}; \cdot, \cdot) : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R}$  is continuously differentiable in the time variable (first unspecified variable) and twice continuously differentiable in the space variable (second unspecified variable).

*Remark 5* Note that Assumption (A2) implies that the diffusion does not explode in finite time (otherwise  $p(s, \mathbf{x}; t, \cdot)$  would determine a sub-probability density). A sufficient condition for non-explosion is to require that  $\mathbf{b}$  is locally Lipschitz in the space variable  $\mathbf{x}$  (which follows from Assumption (A1)) and moreover that there exists a constant  $C$  such that  $|b(t, \mathbf{x})| \leq C(1 + |t| + |\mathbf{x}|)$  for all  $(t, \mathbf{x}) \in [0, \infty) \times \mathbb{R}^d$  [17, Proposition 1.1.11]. Furthermore, the fact that  $\mathbf{b}$  is locally Lipschitz in  $\mathbf{x}$  implies the existence of a unique strong solution to the SDE corresponding to (7) for any given driving Brownian motion  $B$  [17, Theorem 1.1.8].

We will sometimes say  $\mathbf{b}$  satisfies Assumptions (A1) and (A2) if  $\mathbf{b}$  satisfies (A1) and the corresponding diffusion (whose law is unique by the above remark) has transition probability densities satisfying (A2).

Recall that we say a diffusion starts from  $(t, \mathbf{x})$  if we are looking at the law  $\mathcal{L}(\theta_t X | X_t = \mathbf{x})$ , where  $\theta$  denotes the time-shift operator given by  $(\theta_t X)_s = X_{t+s}$ . The resulting process is a diffusion with the identity diffusion matrix but using time-shifted drift  $b(t + \cdot, \cdot)$  and starting from  $x$  at time 0.

Let  $X$  and  $Y$  be two copies of this diffusion starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$  respectively.

Recall

$\mathcal{M}(\mu_s) = \{(\mathbf{x}, \mathbf{y}) \in \text{Support}(\mu_s) : \mathbf{x} \neq \mathbf{y} \text{ and } \mathcal{L}(\theta_s X, \theta_s Y \mid X_s = \mathbf{x}, Y_s = \mathbf{y})$   
yields a maximal coupling of  $(X, Y)$  starting from  $((s, \mathbf{x}), (s, \mathbf{y}))\}$ .

**Remark 6** The function  $(s, \mathbf{x}) \mapsto p(0, \mathbf{x}_0; t - s, \mathbf{x})$  satisfies a backward parabolic equation. Therefore uniqueness theory for such equations yields that there does not exist any  $s > 0$  such that  $p(0, \mathbf{x}_0; s, \mathbf{z}) = p(0, \mathbf{y}_0; s, \mathbf{z})$  for all  $\mathbf{z} \in \mathbb{R}^d$ . This, along with (6), implies that, for every  $s > 0$ ,  $\mu(\tau > s) > 0$  and thus  $\mu(\mathcal{M}(\mu_s)) > 0$ . In particular,  $\mathcal{M}(\mu_s)$  is non-empty for each  $s > 0$ .

## 2.1 Coupling and the interface

Here, we show that the existence of a Markovian maximal coupling for  $X$  and  $Y$  implies that for each time  $t$ , the interface  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  will be a hyperplane bisecting the straight line joining  $X_t$  and  $Y_t$ .

We begin with some preparatory lemmas. Note that Brownian motion has fluctuations which are of order  $O(\sqrt{t})$  while fluctuations resulting from the drift are of order  $O(t)$ . Thus, *on small time scales*, the Brownian behaviour should dominate. The following lemma substantiates this intuition.

**Lemma 7** *Let  $X$  be a diffusion given by*

$$X_t = B_t + \int_0^t \mathbf{b}(s, X_s) \, ds,$$

*with  $X_0 = \mathbf{x}_0$  (so  $B_0 = \mathbf{x}_0$ ), and suppose the drift  $\mathbf{b}$  satisfies Assumption (A1). Denote by  $\mathbb{P}$  the underlying measure. Then, for any  $\mathbf{z} \in \mathbb{R}^d$  and any  $\delta > 0$ ,*

$$\lim_{t \downarrow 0} t \log \frac{\mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} = 0. \quad (8)$$

*Proof* Let  $I = \sup\{|\mathbf{y} - \mathbf{x}_0| : \mathbf{y} \in \mathcal{B}(\mathbf{z}, \delta)\}$  and choose  $N > d \times I + 1$ . By continuity of  $\mathbf{b}$ , there is a finite  $M$  for which  $|\mathbf{b}(t, \mathbf{y})| \leq M$  for all  $(t, \mathbf{y}) \in [0, 1] \times \mathcal{B}(\mathbf{x}_0, N)$ .

Let  $\tau_N = \inf\{t > 0 : X_t \notin \mathcal{B}(\mathbf{x}_0, N)\}$ . Then, we can write

$$\mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta)] = \mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta), \tau_N > t] + \mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta), \tau_N \leq t]. \quad (9)$$

Now  $|X_{t \wedge \tau_N} - B_{t \wedge \tau_N}| \leq Mt$ . We pick  $t \leq \min\{\frac{1}{M}, \frac{\delta}{M}\}$ . Then

$$\frac{\mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} \leq \frac{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta + Mt)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} + \frac{\mathbb{P}[\tau_N \leq t]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} \quad (10)$$

and (using  $t < \delta/M$ )

$$\frac{\mathbb{P}[X_t \in \mathcal{B}(\mathbf{z}, \delta)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} \geq \frac{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta - Mt)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} - \frac{\mathbb{P}[\tau_N \leq t]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]}. \quad (11)$$

Also (using  $t < 1/M$  to control the difference between  $B$  and  $X$ )

$$\begin{aligned}\mathbb{P}[\tau_N \leq t] &\leq \mathbb{P}\left[\sup_{s \leq t} |B_s - \mathbf{x}_0| > N - 1\right] \\ &\leq \frac{4d^2\sqrt{t}}{\sqrt{2\pi}(N-1)} \exp\left(-\frac{(N-1)^2}{2td^2}\right).\end{aligned}$$

Thus, there exists some constant  $C$  such that,

$$\limsup_{t \downarrow 0} t \log \frac{\mathbb{P}(\tau_N \leq t)}{\mathbb{P}(B_t \in \mathcal{B}(\mathbf{z}, \delta))} \leq \limsup_{t \downarrow 0} t \log \left( C \frac{\exp\left(-\frac{(N-1)^2}{2td^2}\right)}{\exp\left(-\frac{t^2}{2t}\right)} \right) < 0. \quad (12)$$

By the Large Deviation principle for Brownian motion [40],

$$\lim_{t \downarrow 0} t \log \frac{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta + Mt)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} = \lim_{t \downarrow 0} t \log \frac{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta - Mt)]}{\mathbb{P}[B_t \in \mathcal{B}(\mathbf{z}, \delta)]} = 0.$$

This, along with (10), (11) and (12), yields the lemma.  $\square$

**Remark 8** The above lemma can be regarded as a weak form of a large deviation principle (LDP) for the diffusion  $X$ , specialized to a particular set  $\mathcal{B}(\mathbf{z}, \delta)$ . The general form of the LDP can be shown to hold under the additional assumption of linear growth of the drift vectorfield, which is used to control the moments of the Radon–Nikodym derivative of the law of  $X$  with respect to that of  $B$  obtained by the Girsanov Theorem [40].

Note that for each fixed  $(s, \mathbf{x})$  the transition density  $(t, \mathbf{y}) \mapsto p(s, \mathbf{x}; t, \mathbf{y})$  satisfies the Kolmogorov forward equation

$$\partial_t p = L^* p \quad (13)$$

where  $L^*$  is the adjoint of the operator  $L$ . Under assumptions (A1) and (A2) the above equation can be rewritten as

$$(\mathcal{A} + h)p = 0,$$

where  $\mathcal{A}$  is a *uniformly parabolic operator* [34, p. 173] and  $h$  is bounded on compact subsets of  $[0, \infty) \times \mathbb{R}^d$ . We now state the *Strong Maximum Principle* for uniformly parabolic equations in the following form (see Theorem 5, Theorem 7 and part (ii) of the remark following Theorem 7, pp. 173–175 of [34]).

**Lemma 9** *Let  $u$  be a solution of*

$$(\mathcal{A} + h)u \geq 0$$

*on a domain of the form  $\Omega_T = (0, T] \times \Omega$ , where  $\Omega$  is a bounded and connected open set and the coefficients of  $\mathcal{A}$ , and the function  $h$  are bounded on closed subsets of  $\Omega_T$ . Suppose  $u \leq 0$  on  $\Omega_T$  and  $u(T, x') = 0$  for some  $x' \in \Omega$ . Then  $u \equiv 0$  on  $\Omega_T$ .*

It is now possible to state and prove the main result of this section, which can be seen as a stronger version of [25, Proposition 3.9], although our proof is quite different and slightly shorter.

**Theorem 10** *Take any  $s > 0$ . For any  $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}(\mu_s)$ , the following equalities hold:*

$$\begin{aligned} I(\mathbf{x}_0, \mathbf{y}_0, s) &= H(\mathbf{x}, \mathbf{y}), \\ I^-(\mathbf{x}_0, \mathbf{y}_0, s) &= H^-(\mathbf{x}, \mathbf{y}), \\ I^+(\mathbf{x}_0, \mathbf{y}_0, s) &= H^+(\mathbf{x}, \mathbf{y}). \end{aligned}$$

*Proof* By continuity of  $\alpha(s, \cdot)$ , it suffices to prove that  $H^-(\mathbf{x}, \mathbf{y}) \subseteq I^-(\mathbf{x}_0, \mathbf{y}_0, s)$  and  $H^+(\mathbf{x}, \mathbf{y}) \subseteq I^+(\mathbf{x}_0, \mathbf{y}_0, s)$ .

We will first show that  $\alpha(s, \mathbf{z}^*) \geq 0$  for all  $\mathbf{z}^* \in H^-(\mathbf{x}, \mathbf{y})$ . Suppose, in contradiction, that  $\alpha(s, \mathbf{z}^*) < 0$  for some  $\mathbf{z}^* \in H^-(\mathbf{x}, \mathbf{y})$ .

Since  $H^-(\mathbf{x}, \mathbf{y})$  is open and  $\alpha$  is continuous, we can choose  $\delta > 0$  such that  $\mathcal{B}(\mathbf{z}^*, \delta) \subseteq H^-(\mathbf{x}, \mathbf{y})$  and  $\alpha(s + s', \mathbf{z}) < 0$  for all  $\mathbf{z} \in \mathcal{B}(\mathbf{z}^*, \delta)$  for sufficiently small  $s' > 0$ . By Lemma 2 this implies that

$$\mu(X_{s+s'} \in \mathcal{B}(\mathbf{z}^*, \delta), \tau > s + s') = 0$$

for all sufficiently small  $s' > 0$ . Let  $B_1, B_2$  be Brownian motions starting from  $\mathbf{x}$  and  $\mathbf{y}$  respectively. Since  $\mathbf{z}^* \in H^-(\mathbf{x}, \mathbf{y})$ , it follows that  $\mathbb{P}[B_{1,t} \in \mathcal{B}(\mathbf{z}^*, \delta)] > \mathbb{P}[B_{2,t} \in \mathcal{B}(\mathbf{z}^*, \delta)]$  for all  $t > 0$ . By Lemma 7, if  $s' > 0$  is sufficiently small then it follows that

$$\mu\left((\theta_s X)_{s'} \in \mathcal{B}(\mathbf{z}^*, \delta) \mid X_s = \mathbf{x}\right) > \mu\left((\theta_s Y)_{s'} \in \mathcal{B}(\mathbf{z}^*, \delta) \mid Y_s = \mathbf{y}\right). \quad (14)$$

By continuity of the transition densities, for all sufficiently small  $s' > 0$  and for small enough open sets  $U_1$  containing  $\mathbf{x}$  and  $U_2$  containing  $\mathbf{y}$ , for any  $(\mathbf{u}_1, \mathbf{u}_2) \in (U_1 \times U_2) \cap \mathcal{M}(\mu_s)$ ,

$$\begin{aligned} &\mu\left(X_{s+s'} \in \mathcal{B}(\mathbf{z}^*, \delta), \tau > s + s' \mid X_s = \mathbf{u}_1, Y_s = \mathbf{u}_2\right) \\ &= \int_{\mathcal{B}(\mathbf{z}^*, \delta)} \alpha^+(s, \mathbf{u}_1, \mathbf{u}_2, s + s', \mathbf{z}) \, d\mathbf{z} \\ &\geq \int_{\mathcal{B}(\mathbf{z}^*, \delta)} \alpha(s, \mathbf{u}_1, \mathbf{u}_2, s + s', \mathbf{z}) \, d\mathbf{z} \\ &= \mu\left((\theta_s X)_{s'} \in \mathcal{B}(\mathbf{z}^*, \delta) \mid X_s = \mathbf{u}_1\right) - \mu\left((\theta_s Y)_{s'} \in \mathcal{B}(\mathbf{z}^*, \delta) \mid Y_s = \mathbf{u}_2\right) \\ &> 0. \end{aligned} \quad (15)$$

(Here, the first equality follows from Lemmas 2 and 3.) Since  $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}(\mu_s)$ , it follows that  $\mu((X_s, Y_s) \in (U_1 \times U_2) \cap \mathcal{M}(\mu_s)) > 0$ , yielding (for all sufficiently small  $s > 0$ )

$$\mu(X_{s+s'} \in \mathcal{B}(\mathbf{z}^*, \delta), \tau > s + s') > 0,$$

contradicting our assumption. Hence  $\alpha(s, \mathbf{z}^*) \geq 0$  for all  $\mathbf{z}^* \in H^-(\mathbf{x}, \mathbf{y})$ . Similarly,  $\alpha(s, \mathbf{z}^*) \leq 0$  for all  $\mathbf{z}^* \in H^+(\mathbf{x}, \mathbf{y})$ .

We have thus shown that

$$\begin{aligned} H^-(\mathbf{x}, \mathbf{y}) &\subseteq I(\mathbf{x}_0, \mathbf{y}_0, s) \cup I^-(\mathbf{x}_0, \mathbf{y}_0, s), \\ H^+(\mathbf{x}, \mathbf{y}) &\subseteq I(\mathbf{x}_0, \mathbf{y}_0, s) \cup I^+(\mathbf{x}_0, \mathbf{y}_0, s). \end{aligned}$$

Suppose  $H^-(\mathbf{x}, \mathbf{y}) \cap I(\mathbf{x}_0, \mathbf{y}_0, s)$  is non-empty, and pick  $\mathbf{z}^* \in H^-(\mathbf{x}, \mathbf{y}) \cap I(\mathbf{x}_0, \mathbf{y}_0, s)$ . Since  $\alpha(s, \cdot)$  is nonnegative on the open set  $H^-(\mathbf{x}, \mathbf{y})$ , there exists  $\delta > 0$  such that  $\alpha(s, \mathbf{z}) \geq 0$  for all  $\mathbf{z} \in \mathcal{B}(\mathbf{z}^*, \delta)$ . Choose open sets  $U_1$  containing  $\mathbf{x}$  and  $U_2$  containing  $\mathbf{y}$ , and possibly smaller  $\delta > 0$ , such that  $|\mathbf{x}' - \mathbf{z}| < |\mathbf{y}' - \mathbf{z}|$  for all  $\mathbf{x}' \in U_1, \mathbf{y}' \in U_2$  and  $\mathbf{z} \in \mathcal{B}(\mathbf{z}^*, \delta)$ . It is given that  $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}(\mu_s)$ ; since the process  $((X_t, Y_t) : t \geq 0)$  has continuous paths there must be  $\eta > 0$  such that  $\mu_t(U_1 \times U_2) > 0$  for all  $t \in [s - \eta, s]$ .

The function  $(t, \mathbf{z}) \mapsto \alpha(t, \mathbf{z})$  solves the Kolmogorov forwards Eq. (13). Thus we can apply Lemma 9 to  $-\alpha$  on  $\Omega_\eta = (s - \eta, s] \times \mathcal{B}(\mathbf{z}^*, \delta)$ , and deduce that either  $\alpha(t, \mathbf{z}) = 0$  for all  $s - \eta < t < s$  and all  $\mathbf{z} \in \mathcal{B}(\mathbf{z}^*, \delta)$ , or there exists  $s' \in (s - \eta, s)$ ,  $0 < \varepsilon < s - s'$  and an open set  $U \subseteq \mathcal{B}(\mathbf{z}^*, \delta)$  such that  $\alpha(t, \mathbf{z}) < 0$  for all  $\mathbf{z} \in U$  and all  $t \in [s', s' + \varepsilon]$ . In either case (taking  $U = \mathcal{B}(\mathbf{z}^*, \delta)$  and any  $s' \in (s - \eta, s)$ ,  $\varepsilon \in (0, s - s')$  in the first case), for all  $t \in [s', s' + \varepsilon]$

$$\mu(X_t \in U, \tau > t) = 0. \quad (16)$$

Now choose  $(\mathbf{x}', \mathbf{y}') \in (U_1 \times U_2) \cap \mathcal{M}(\mu_{s'})$  (non-empty, since  $U_1$  and  $U_2$  are disjoint and  $\mu_{s'}(U_1 \times U_2) > 0$ ) and apply the same argument as the one used in obtaining (15), but with  $\mathbf{x}', \mathbf{y}'$  replacing  $\mathbf{x}, \mathbf{y}$  and  $s'$  replacing  $s$ . We obtain

$$\mu(X_{s'+s''} \in U, \tau > s' + s'') > 0$$

for some  $s'' \in [s', s' + \varepsilon]$ , contradicting (16). The lemma follows.  $\square$

**Remark 11** The above theorem shows that for a Markovian maximal coupling, for any time  $s$ , the locus  $I(\mathbf{x}_0, \mathbf{y}_0, s)$  can be viewed as a (possibly time-varying) *mirror* which realizes the coupling in a very explicit way, using a (possibly time-varying) reflection isometry.

The following corollary to the above lemma shows that the coupling time  $\tau$  is, in fact, the hitting time of the deterministic space-time set  $\{(s, I(\mathbf{x}_0, \mathbf{y}_0, s)) : s > 0\}$  by the process  $((s, X_s) : s \geq 0)$  (equivalently,  $((s, Y_s) : s > 0)$ ). In particular,  $X$  and  $Y$  will couple at the first time they meet. Furthermore, the interface representation described in Theorem 10 will hold almost surely for all time before coupling occurs.

**Corollary 12** Consider a Markovian maximal coupling, with coupling time  $\tau$ . Set  $\tau' = \inf\{s > 0 : X_s \in I(\mathbf{x}_0, \mathbf{y}_0, s)\}$ . Almost surely  $\tau = \tau'$ . Furthermore,  $\mu$ -almost surely, for all  $t < \tau$ ,

$$\begin{aligned} I(\mathbf{x}_0, \mathbf{y}_0, t) &= H(X_t, Y_t), \\ I^-(\mathbf{x}_0, \mathbf{y}_0, t) &= H^-(X_t, Y_t), \\ I^+(\mathbf{x}_0, \mathbf{y}_0, t) &= H^+(X_t, Y_t). \end{aligned} \quad (17)$$

*Proof* Note that, by Lemma 2,

$$\mu(Y_q \in I^-(\mathbf{x}_0, \mathbf{y}_0, q) \text{ for some rational } q < \tau) = 0.$$

Since the trajectories of  $Y$  are continuous, it follows that almost surely  $Y_t$  is contained in the complement of  $I^-(\mathbf{x}_0, \mathbf{y}_0, t)$  for all  $t < \tau$ . This implies that before time  $\tau'$ ,  $X$  and  $Y$  are supported on disjoint subsets of the state space and hence

$$\mu(\tau' \leq \tau) = 1. \quad (18)$$

For any  $t > 0$ , we define the event

$$E_t = \left[ \text{Either } X_t = Y_t, \text{ or } X_t \neq Y_t \text{ and all three equalities } I(\mathbf{x}_0, \mathbf{y}_0, t) = H(X_t, Y_t), \right. \\ \left. I^-(\mathbf{x}_0, \mathbf{y}_0, t) = H^-(X_t, Y_t), \text{ } I^+(\mathbf{x}_0, \mathbf{y}_0, t) = H^+(X_t, Y_t) \text{ hold.} \right]. \quad (19)$$

Theorem 10 implies the assertion

$$\mu(E_q \text{ is true for all rational } q) = 1, \quad (20)$$

hence almost surely  $E = \bigcap_{q \in \mathbb{Q}} E_q$  holds. Take any  $t > 0$  with  $X_t \neq Y_t$  and let  $\mathbf{z} \in H(X_t, Y_t)$ . Then it follows from the definition of  $H(\mathbf{x}, \mathbf{y})$  and the continuity of sample paths of  $X$  and  $Y$  that there is a rational sequence  $t_n \downarrow t$  and  $\mathbf{z}_n \in H(X_{t_n}, Y_{t_n})$  such that  $\mathbf{z}_n \rightarrow \mathbf{z}$ . Thus, on the event  $E$ , the continuity of  $\alpha$  implies that  $H(X_t, Y_t) \subseteq I(\mathbf{x}_0, \mathbf{y}_0, t)$ .

Now, take  $\mathbf{z} \in H^+(X_t, Y_t)$  when  $X_t \neq Y_t$ . The continuity of sample paths of  $X$  and  $Y$  implies that there exist  $\eta, \delta > 0$  with  $\mathcal{B}(\mathbf{z}, \eta) \subseteq H^+(X_s, Y_s)$  for all  $s \in [t - \delta, t]$ . On the event  $E$ , the continuity of  $\alpha$  implies  $\alpha(s, \mathbf{z}') \leq 0$  for all  $s \in [t - \delta, t]$  when  $\mathbf{z}' \in \mathcal{B}(\mathbf{z}, \eta)$ . Thus, as  $\alpha(q, \mathbf{z}) < 0$  for all rational  $q \in [t - \delta, t]$ , Lemma 9 implies  $\alpha(t, \mathbf{z}) < 0$ . Thus,  $H^+(X_t, Y_t) \subseteq I^+(\mathbf{x}_0, \mathbf{y}_0, t)$ . Similarly,  $H^-(X_t, Y_t) \subseteq I^-(\mathbf{x}_0, \mathbf{y}_0, t)$ . As  $\mu(E) = 1$ , it follows that

$$\mu(E_t \text{ is true for all } t) = 1. \quad (21)$$

Note that, in particular, (18) and (21) imply that if  $\tau' < \infty$ , then  $X_{\tau'} = Y_{\tau'}$  almost surely. For  $\|X_t - Y_t\| = \frac{1}{2} \text{dist}(X_t, H(X_t, Y_t)) = \frac{1}{2} \text{dist}(X_t, I(\mathbf{x}_0, \mathbf{y}_0, t))$  (when  $t < \tau'$ ), by definition of  $H(X_t, Y_t)$ .

The corresponding argument for  $Y$  implies that  $\tau'$  also satisfies  $\tau' = \inf\{s > 0 : Y_s \in I(\mathbf{x}_0, \mathbf{y}_0, s)\}$ . Therefore,  $\tau'$  is a stopping time for both  $X$  and  $Y$ . Since  $X_{\tau'} = Y_{\tau'}$ , we can extend  $X$  and  $Y$  synchronously beyond time  $\tau'$ . Combined with

(18), this implies  $\tau = \tau'$  almost surely, since the maximal coupling time  $\tau$  must be stochastically smaller than all other coupling times. Consequently

$$\mu(X_t \neq Y_t \text{ for all } t < \tau) = 1.$$

This, together with (21), yields (17) and thus the corollary is proved.  $\square$

## 2.2 Time evolution of the mirror

We now analyze the time-evolution of the mirror. From Theorem 10, it follows that the mirror  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  is a hyperplane for each  $t > 0$ . We parametrize this hyperplane by its signed distance from the origin, say  $l(t)$ , together with the normal vector to the hyperplane, say  $\mathbf{n}(t)$ . There is an ambiguity of sign in the choice of  $\mathbf{n}(t)$ ; however the next lemma states that  $\mathbf{n}(t)$  can be chosen to make this parametrization continuous up to the coupling time  $\tau$ .

**Lemma 13** *Suppose that a Markovian maximal coupling exists for  $X$  and  $Y$ . Then there exists a continuous parametrization  $((l(t), \mathbf{n}(t)) : t \in [0, \tau))$  of  $I(\mathbf{x}_0, \mathbf{y}_0, \cdot)$ .*

*Proof* Corollary 12, together with the remark following Lemma 3, shows that the following subset of coupled path-space  $C[0, \infty)^2$  is non-empty for any  $S > 0$ , and indeed of full  $\mu$ -measure in the subset corresponding to  $\tau > S$ :

$$A_S = \{\omega \in C[0, \infty)^2 : I(\mathbf{x}_0, \mathbf{y}_0, t) = H(X_t(\omega), Y_t(\omega)) \text{ for all } t \leq S, \tau > S\}.$$

Consider any coupled pair of paths  $\omega \in A_S$ . Define  $(l(t), \mathbf{n}(t))$  on  $[0, S]$  by

$$\begin{aligned} \mathbf{n}^{(S)}(t) &= \frac{X_t(\omega) - Y_t(\omega)}{|X_t(\omega) - Y_t(\omega)|}, \\ l^{(S)}(t) &= \mathbf{n}^\top(t) \left( \frac{X_t(\omega) + Y_t(\omega)}{2} \right). \end{aligned} \quad (22)$$

This gives a continuous parametrization  $(l^{(S)}, \mathbf{n}^{(S)})$  on  $[0, S \wedge \tau)$ .

This recipe can be used to define  $(l^{(N)}, \mathbf{n}^{(N)})$  on  $[0, N \wedge \tau)$  for each positive integer  $N$ . By continuity of  $\mathbf{n}^{(N)}$  and  $\mathbf{n}^{(N+1)}$  on the (connected) interval  $[0, N \wedge \tau)$ , we see that either  $\mathbf{n}^{(N)} \equiv \mathbf{n}^{(N+1)}$  or  $\mathbf{n}^{(N)} \equiv -\mathbf{n}^{(N+1)}$  on  $[0, N \wedge \tau)$ . But

$$\lim_{t \downarrow 0} \mathbf{n}^{(N)}(t) = \lim_{t \downarrow 0} \mathbf{n}^{(N+1)}(t) = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|},$$

implying  $\mathbf{n}^{(N)} \equiv \mathbf{n}^{(N+1)}$  on  $[0, N \wedge \tau)$ . Consequently  $l^{(N)} = l^{(N+1)}$  on  $[0, N \wedge \tau)$ . So we can consistently and continuously define the parametrization as  $((l(t), \mathbf{n}(t)) : t \in [0, \tau))$ , thus proving the lemma.  $\square$

In fact the parametrization is not simply continuous but is also continuously differentiable:

**Lemma 14** *Suppose that a Markovian maximal coupling exists for  $X$  and  $Y$ . Then the parametrization  $(l(t), \mathbf{n}(t))$  of the mirror  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  (defined for  $t \in [0, \tau)$ ) is continuously differentiable in  $t$ .*

*Proof* We use the fact that the map given by reflection in the hyperplane parametrized by  $(l(t), \mathbf{n}(t))$ ,

$$F(t, \mathbf{x}) = (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))\mathbf{x} + 2l(t)\mathbf{n}(t),$$

takes  $X_t$  to  $Y_t$  for  $t \in [0, \tau)$  (this follows from  $I(\mathbf{x}_0, \mathbf{y}_0, t) = H(X_t, Y_t)$ ). Take any  $\mathbf{x} \in I^-(\mathbf{x}_0, \mathbf{y}_0, t)$ . Let  $U$  be an open ball containing  $\mathbf{x}$  and contained in  $I^-(\mathbf{x}_0, \mathbf{y}_0, t)$ . Let  $\tau_U = \inf\{s > t : X_s \notin U\}$ . Consider the corresponding stopped processes  $X_s^U = X_{s \wedge \tau_U}$  and  $Y_s^U = Y_{s \wedge \tau_U}$  for  $s \geq t$ . We write expectation with respect to  $\mu$  using  $\mathbb{E}$ .

By general properties of diffusions [31, Chapter 11],

$$\begin{aligned} \mathbf{b}(t, \mathbf{x}) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{X_s^U - \mathbf{x}}{s - t} \mid X_t^U = \mathbf{x} \right], \\ \mathbf{b}(t, F(t, \mathbf{x})) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{Y_s^U - F(t, \mathbf{x})}{s - t} \mid Y_t^U = F(t, \mathbf{x}) \right]. \end{aligned} \quad (23)$$

Note that under the coupling  $\mu$  we may use Corollary 12 to see that  $Y_s^U = F(s, X_s^U)$  for all  $s \geq t$  with probability one. Thus, we can write the last expression above as

$$\begin{aligned} \mathbf{b}(t, F(t, \mathbf{x})) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{F(s, X_s^U) - F(t, \mathbf{x})}{s - t} \mid X_t^U = \mathbf{x} \right] \\ &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{F(s, X_s^U) - F(s, \mathbf{x})}{s - t} \mid X_t^U = \mathbf{x} \right] + \lim_{s \downarrow t} \frac{F(s, \mathbf{x}) - F(t, \mathbf{x})}{s - t}, \end{aligned}$$

in the sense that if the limit of  $\mathbb{E} \left[ \frac{F(s, X_s^U) - F(s, \mathbf{x})}{s - t} \mid X_t^U = \mathbf{x} \right]$  exists then also the limit of  $\frac{F(s, \mathbf{x}) - F(t, \mathbf{x})}{s - t}$  exists and is defined by the above. By linearity of  $F$  in  $\mathbf{x}$ , we see that the first summand becomes

$$\begin{aligned} &\lim_{s \downarrow t} \mathbb{E} \left[ \frac{F(s, X_s^U) - F(s, \mathbf{x})}{s - t} \mid X_t^U = \mathbf{x} \right] \\ &= (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t)) \lim_{s \downarrow t} \mathbb{E} \left[ \frac{X_s^U - \mathbf{x}}{s - t} \mid X_t^U = \mathbf{x} \right] \\ &= (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))\mathbf{b}(t, \mathbf{x}). \end{aligned}$$

This shows that  $\lim_{s \downarrow t} \frac{F(s, \mathbf{x}) - F(t, \mathbf{x})}{s - t}$  exists for each  $\mathbf{x}$  and for all  $t \in [0, \tau)$  and indeed is continuous in  $t$ . This is enough to show that  $t \mapsto F(t, \mathbf{x})$  is continuously differentiable for each  $\mathbf{x}$  [4, Theorem 1.3]. This follows from the facts that  $t \mapsto (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))$  and  $t \mapsto l(t)\mathbf{n}(t)$  are continuously differentiable, and actually requires these facts to be true: consider  $F(t, \mathbf{x})$  for  $\mathbf{x}$  varying over an orthonormal basis and also for  $\mathbf{x} = 0$ .



Now, take any  $t_0 \in [0, \tau)$ . Let  $n_i$  denote the  $i$ th component of  $\mathbf{n}$ . As  $|\mathbf{n}(t_0)| = 1$ , there is an  $i$  such that  $n_i(t) \neq 0$  in a neighbourhood  $V$  of  $t_0$ . The continuous differentiability of  $t \mapsto (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))$  implies  $n_i n_j$  is continuously differentiable in  $V$  for all  $1 \leq j \leq d$ . This implies  $n_j$  is continuously differentiable in  $V$  for all  $j$ . Differentiability of  $t \mapsto l(t)\mathbf{n}(t)$  then shows that  $l$  is continuously differentiable on  $V$ . This proves the lemma.  $\square$

## 2.3 Structure of the coupling

All the tools having been assembled, it is now possible to present a rather explicit description of drifts  $\mathbf{b}$  which permit the existence of a Markovian maximal coupling of two copies  $X$  and  $Y$  of a Euclidean diffusion with the required regularity conditions.

We begin with a notational remark. For any  $\mathbf{x} \in \mathbb{R}^d$  and any hyperplane  $\underline{h}$ , we denote by  $\underline{h}\mathbf{x}$  the reflection of  $\mathbf{x}$  in  $\underline{h}$ . We write  $\underline{h}_k$  for the hyperplane  $\{x_k = 0\}$ .

The first lemma of this subsection concerns an observation concerning rotations and shifts of these Euclidean diffusions.

**Lemma 15** *Let  $X$  be an Euclidean diffusion satisfying assumptions (A1), (A2). Let  $Q : [0, \infty) \rightarrow \mathbf{O}(d)$  be a continuously differentiable function taking values in the space of orthogonal  $(d \times d)$  matrices, and let  $l : [0, \infty) \rightarrow \mathbb{R}$  be a continuously differentiable real-valued function. Then the new process given by*

$$\tilde{X}_t = Q(t)X_t - l(t)\mathbf{e}_1 \quad (24)$$

*satisfies the stochastic differential equation*

$$d\tilde{X}_t = \tilde{\mathbf{b}}(t, \tilde{X}_t) dt + d\tilde{B}_t \quad (25)$$

*where*

$$\tilde{\mathbf{b}}(t, x) = \dot{Q}(t)Q^T(t)(x + l(t)\mathbf{e}_1) + Q(t)\mathbf{b}(t, Q^T(t)(x + l(t)\mathbf{e}_1)) - \dot{l}(t)\mathbf{e}_1 \quad (26)$$

*and*

$$d\tilde{B}_t = Q(t) dB_t. \quad (27)$$

*Here,  $\dot{Q}$  and  $\dot{l}$  denote the respective time-derivatives and  $Q^\top$  denotes the matrix transpose.*

*Proof* The result follows by direct calculation using Itô calculus.  $\square$

**Remark 16** Note that the transformed drift given by (26) satisfies the regularity Assumptions (A1) and (A2). (A1) follows via the explicit form of (26) from the fact that  $\mathbf{b}$  satisfies (A1) and  $Q$  and  $l$  are continuously differentiable. (A2) for the new process  $\tilde{X}$  follows from (24) and the fact that  $X$  satisfies (A2).

The following theorem describes Markovian maximal couplings for the class of time-nonhomogeneous Euclidean diffusions satisfying suitable regularity conditions. The intuitive content of the theorem is, given an MMC  $(X, Y)$ , *applying deterministic time-varying rotations and translations to the ambient Euclidean space reduces this MMC to a reflection coupling in a fixed hyperplane*. Thus, in a certain sense, reflection coupling is the only type of Markovian coupling that can possibly preserve maximality.

**Theorem 17** *Let  $X$  be an Euclidean diffusion starting from  $\mathbf{x}_0$  and satisfying assumptions (A1), (A2).*

- (i) *Suppose the following holds for every  $x \in \mathbb{R}^d$ , for the fixed hyperplane  $\underline{h}_1 = \{x_1 = 0\}$ .*

$$\mathbf{b}(t, \underline{h}_1 \mathbf{x}) = \underline{h}_1 \mathbf{b}(t, \mathbf{x}) \quad (28)$$

*Then, for  $\tau_0 = \inf\{t \geq 0 : X_t \in \underline{h}_1\}$ , the reflection-coupling*

$$Y_t = \begin{cases} \underline{h}_1 X_t & \text{if } t < \tau_0 \\ X_t & \text{if } t \geq \tau_0 \end{cases} \quad (29)$$

*gives a Markovian maximal coupling between two copies of the diffusion starting from  $\mathbf{x}_0$  and  $\underline{h}_1 \mathbf{x}_0$  respectively.*

- (ii) *Let  $Y$  be a coupled copy of  $X$ . Then  $(X, Y)$  is a Markovian maximal coupling up to the maximal coupling time  $\tau$  if and only if there exist  $C^1$  curves  $Q : [0, \tau) \rightarrow \mathbf{O}(d)$  and  $l : [0, \tau) \rightarrow \mathbb{R}$  (compare Lemma 15) with  $Q(0) \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|} = \mathbf{e}_1$  and  $l(0) = \frac{|\mathbf{x}_0|^2 - |\mathbf{y}_0|^2}{2|\mathbf{x}_0 - \mathbf{y}_0|}$ , such that  $(\tilde{X}, \tilde{Y})$  obtained from  $(X, Y)$  using the transformation (24) are reflection-coupled according to the recipe (29). In particular, the transformed time-varying drift  $\tilde{\mathbf{b}}$  given by (26) must satisfy*

$$\tilde{\mathbf{b}}(t, \underline{h}_1 \mathbf{x}) = \underline{h}_1 \tilde{\mathbf{b}}(t, \mathbf{x}). \quad (30)$$

- Proof* (i) Equation (28) implies that the process  $(\underline{h}_1 X_t : t \geq 0)$  has the same law as the diffusion starting from  $\underline{h}_1 \mathbf{x}_0$  and thus, the reflection-coupling (29) gives a valid coupling. Reflection in the hyperplane  $\underline{h}_1$  thus gives a *reflection structure* in the sense of [24, Definition 2.1]. Maximality follows from [24, Proposition 2.2].
- (ii) First, note that if  $\tilde{X}$  and  $\tilde{Y}$  are reflection-coupled according to (29), then analysis of generators of  $\underline{h}_1 \tilde{X}_t$  and  $\tilde{Y}_t$  yields (30). Now, applying part (i) of the theorem, we deduce that  $(\tilde{X}, \tilde{Y})$  is a Markovian maximal coupling. Furthermore, as

$$(t, x) \mapsto (t, Q^\top(t)(x + l(t)\mathbf{e}_1))$$

is a bijective, bimeasurable function, so application of Lemma 4 to  $(t, \tilde{X}_t) \rightarrow (t, X_t)$  and  $(t, \tilde{Y}_t) \rightarrow (t, Y_t)$  shows that  $(X, Y)$  is a Markovian maximal coupling. Conversely, let  $(X, Y)$  be a Markovian maximal coupling of two copies of the diffusion starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . Then the results of Sects. 2.1 and 2.2 show that there exist continuously differentiable functions  $l : [0, \infty) \rightarrow \mathbb{R}$  and  $\mathbf{n} :$

$[0, \infty) \rightarrow \mathbb{S}^{d-1}$  parametrising the mirror  $I(\mathbf{x}_0, \mathbf{y}_0, t)$ . Moreover, these functions should satisfy  $\mathbf{n}(0) = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  and  $l(0) = \frac{|\mathbf{x}_0|^2 - |\mathbf{y}_0|^2}{|\mathbf{x}_0 - \mathbf{y}_0|}$ . To see this, take  $t \downarrow 0$  in (22). Furthermore, Theorem 10 and the corollary following it show that  $X$  and  $Y$  are coupled on  $t < \tau$  according to the relationship

$$Y_t = (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))X_t + 2l(t)\mathbf{n}(t). \quad (31)$$

The construction of  $Q$  follows by applying Gram–Schmidt orthogonalization to extend  $\mathbf{n}(0)$  to an orthonormal basis  $(\mathbf{n}(0), \mathbf{v}_1, \dots, \mathbf{v}_{d-1})$  of  $\mathbb{R}^d$ . Note that the vectors  $\mathbf{v}_i$  lie in the tangent space of  $\mathbb{S}^{d-1}$  based at  $\mathbf{n}(0)$ . The vector function  $(\mathbf{n}(t) : t \geq 0)$  traces out a  $C^1$  curve  $\gamma$  on the sphere  $\mathbb{S}^{d-1}$ . Parallel transport [14, p. 75] can be applied along  $\gamma$  to each vector  $\mathbf{v}_i$ ; this produces  $C^1$  vectorfields  $\mathbf{X}_i : [0, \infty) \rightarrow \mathbb{R}^d$  along  $\gamma$ . [14, Proposition 2.74] shows that  $(\mathbf{n}, \mathbf{X}_1, \dots, \mathbf{X}_{d-1})$  produces a  $C^1$  orthonormal frame along  $\gamma$ , so set

$$Q^\top(t) = (\mathbf{n}(t), \mathbf{X}_1(t), \dots, \mathbf{X}_{d-1}(t)).$$

We now produce a new pair of diffusions with time-varying drifts,  $(\tilde{X}, \tilde{Y})$ , by applying the transformation (24) to  $(X, Y)$  with drift  $\tilde{\mathbf{b}}$  and driving Brownian motion  $\tilde{B}$  as described in Lemma 15. This new pair is also a Markovian maximal coupling (use Lemma 4), and from Eq. (31) it follows that the coupled pair  $(\tilde{X}, \tilde{Y})$  is described by the transformation (29). As discussed in part (i) of this proof, the relationship (30) follows as a direct consequence.  $\square$

Inverting the relationship (26), and using the relationship (30), the above theorem yields the following characterisation of drifts which permit MMC:

**Corollary 18** *Under assumptions (A1) and (A2), the Markovian coupling of  $d$ -dimensional Euclidean diffusions  $(X, Y)$  is a Markovian maximal coupling if and only if there exist function  $Q : [0, \tau) \rightarrow \mathbf{O}(d)$  and  $l : [0, \tau) \rightarrow \mathbb{R}$ , as prescribed in Theorem 17, such that*

$$\mathbf{b}(t, \mathbf{x}) = Q^\top(t)\tilde{\mathbf{b}}(t, Q(t)\mathbf{x} - l(t)\mathbf{e}_1) - Q^\top(t)\dot{Q}(t)\mathbf{x} + \dot{l}(t)\mathbf{n}(t) \quad (32)$$

for some  $\tilde{\mathbf{b}}$  satisfying Assumptions (A1) and (A2) and fulfilling the relationship (30).

## 2.4 Rigidity theorems for time-homogeneous diffusions

The previous subsection established an implicit classification of all time-nonhomogeneous diffusions that can be coupled by a Markovian maximal coupling. But, as noted in the literature, not many examples of such couplings are known for *time-homogeneous* diffusions. It is a matter of general belief that the class of such time-homogeneous diffusions is very small, but little rigorous work appears to have been done to specify this class.

In this subsection we obtain a constraint equation on the drift, leading to certain general conditions on the drift and the starting points which are necessary for the

existence of Markovian maximal couplings. In the case of affine drifts the constraint equations are explicit enough to classify all affine drifts leading to Markovian maximal couplings. We then state and prove the main theorem of this subsection: if there are two balls  $\mathcal{B}(\mathbf{x}_0, r)$  and  $\mathcal{B}(\mathbf{y}_0, r)$  in  $\mathbb{R}^d$ , such that a Markovian maximal coupling exists from all pairs of points  $(\mathbf{x}, \mathbf{y}) \in \mathcal{B}(\mathbf{x}_0, r) \times \mathcal{B}(\mathbf{y}_0, r)$ , then the drift has to be of a very simple affine form, verifying the popular belief that Markovian maximal couplings are indeed very rare.

We conclude by showing a stronger result for one-dimensional diffusions, which states that for such a coupling to exist for a specific pair of starting points, either the drift must be an odd function centred at a point, or it must be affine.

The following lemma supplies the constraint equation on the drift. Recall that

$$F(t, \mathbf{x}) = (\mathbb{I} - 2\mathbf{n}(t)\mathbf{n}^\top(t))\mathbf{x} + 2l(t)\mathbf{n}(t) \quad (33)$$

is a linear transformation sending  $\mathbf{x} \in \mathbb{R}^d$  to its reflection in the mirror  $I(\mathbf{x}_0, \mathbf{y}_0, t)$ . For the sake of concise exposition, in the following two lemmas and their proofs we suppress the argument  $t$  when writing  $l$  and  $\mathbf{n}$ .

**Lemma 19** *Assume (A1), (A2) hold. A Markovian maximal coupling  $(X, Y)$  exists from starting points  $\mathbf{x}_0$  and  $\mathbf{y}_0$  if and only if there exist continuously differentiable functions  $l : [0, \infty) \rightarrow \mathbb{R}$  and  $\mathbf{n} : [0, \infty) \rightarrow \mathbb{S}^{d-1}$ , with  $\mathbf{n}(0) = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  and  $l(0) = \frac{|\mathbf{x}_0|^2 - |\mathbf{y}_0|^2}{2|\mathbf{x}_0 - \mathbf{y}_0|}$ , for which the drift vectorfield  $\mathbf{b}$  satisfies the following equation:*

$$\mathbf{b}(\mathbf{x}) = 2(\dot{\mathbf{n}}\mathbf{n}^\top - \mathbf{n}\dot{\mathbf{n}}^\top)\mathbf{x} + 2(\dot{l}\mathbf{n} - l\dot{\mathbf{n}}) + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{b}(F(t, \mathbf{x})). \quad (34)$$

*Proof* First, assume that a Markovian maximal coupling  $(X, Y)$  exists. Note from Eq. (31) that

$$Y_t = F(t, X_t)$$

for  $t \in [0, \tau)$ , with  $\{(l(t), \mathbf{n}(t)) : t \in [0, \tau)\}$  obtained from Lemmas 13 and 14. Applying stochastic calculus to the function  $F$  for  $t \in [0, \tau)$ , substituting in

$$X_t = (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)(Y_t - 2l\mathbf{n}),$$

and simplifying, we obtain

$$\begin{aligned} dY_t &= \left( 2(\dot{\mathbf{n}}\mathbf{n}^\top - \mathbf{n}\dot{\mathbf{n}}^\top)Y_t + 2(\dot{l}\mathbf{n} - l\dot{\mathbf{n}}) + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{b}(F(t, Y_t)) \right) \\ &\quad \times dt + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)dB_t. \end{aligned} \quad (35)$$

The diffusion term is clearly a Brownian motion, as can be verified by the Lévy criterion. On the other hand, the drift term in the semimartingale decomposition of  $Y$  is given by  $\mathbf{b}(t, Y_t)dt$ . Equating the two drifts yields the necessity of the drift constraint condition (34).

Now, suppose  $\mathbf{b}$  satisfies (34) for  $l$  and  $\mathbf{n}$  as given in the lemma. Let  $\tau = \inf\{t > 0 : X_t \in I(\mathbf{x}_0, \mathbf{y}_0, t)\}$ . Then (35) shows that  $Y_t = F(t, X_t)\mathbb{I}(t < \tau) + X_t\mathbb{I}(t \geq \tau)$  gives a valid coupling  $\mu$  of the two copies  $(X, Y)$  with coupling time  $\tau$ . To see that this is indeed the maximal coupling, obtain the  $C^1$  curve  $Q : [0, \tau] \rightarrow \mathbf{O}(d)$  from  $\mathbf{n}$  by the procedure given in the proof of Theorem 17 (ii). Now,  $(\tilde{X}, \tilde{Y})$  obtained from  $(X, Y)$  by (24) is reflection-coupled according to the recipe in (29). Theorem 17 (ii) then implies that  $(X, Y)$  is a Markovian maximal coupling.  $\square$

Equation (34) provides the constraint only in implicit form, and the main task is to extract as much information from it as possible. In what follows, we decompose the gradient matrix  $\nabla \mathbf{b}$  into symmetric and skew-symmetric parts *via*

$$\nabla \mathbf{b}(\mathbf{x}) = S(\mathbf{x}) + T(\mathbf{x}), \quad (36)$$

where  $S(\mathbf{x}) = \frac{\nabla \mathbf{b}(\mathbf{x}) + (\nabla \mathbf{b})^\top(\mathbf{x})}{2}$  and  $T(\mathbf{x}) = \frac{\nabla \mathbf{b}(\mathbf{x}) - (\nabla \mathbf{b})^\top(\mathbf{x})}{2}$ . The next lemma records relations for  $S(\mathbf{x})$  and  $T(\mathbf{x})$  which are direct consequences of (34).

**Lemma 20** *Under the hypotheses of Lemma 19 and (34), the following hold for all  $\mathbf{x} \in \mathbb{R}^d$  and  $t > 0$ :*

(i)

$$S(\mathbf{x}) = (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)S(F(t, \mathbf{x}))(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top), \quad (37)$$

and

$$T(\mathbf{x}) = 2(\dot{\mathbf{n}}\mathbf{n}^\top - \mathbf{n}\dot{\mathbf{n}}^\top) + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)T(F(t, \mathbf{x}))(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top). \quad (38)$$

*In particular,  $S(\mathbf{x})$  and  $S(F(t, \mathbf{x}))$  have the same set of eigenvalues.*

(ii) *There exists a continuous function  $\lambda(\cdot, \cdot) : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that*

$$\left( \frac{S(\mathbf{x}) + S(F(t, \mathbf{x}))}{2} \right) \mathbf{n} = \lambda(t, \mathbf{x})\mathbf{n}. \quad (39)$$

(iii)

$$\left( \frac{T(\mathbf{x}) + T(F(t, \mathbf{x}))}{2} \right) \mathbf{n} = \dot{\mathbf{n}}. \quad (40)$$

*Proof* Differentiating both sides of (34), while recalling the reflection form of  $F(t, \mathbf{x})$  as given in (33), we obtain

$$\nabla \mathbf{b}(\mathbf{x}) = 2(\dot{\mathbf{n}}\mathbf{n}^\top - \mathbf{n}\dot{\mathbf{n}}^\top) + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\nabla \mathbf{b}(F(t, \mathbf{x}))(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top). \quad (41)$$

This immediately yields part (i). The equality of the set of eigenvalues follows from the fact that the reflection matrix  $(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)$  is symmetric and orthogonal.

Parts (ii) and (iii) follow by post-multiplying the equations of part (i) by  $\mathbf{n}$ , bearing in mind that as  $\mathbf{n}$  is a unit vector therefore  $\mathbf{n}$  and  $\dot{\mathbf{n}}$  must be orthogonal.  $\square$

Because  $\mathbf{n}(0) = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  and  $l(0) = \mathbf{n}(0) \cdot \frac{\mathbf{x}_0 + \mathbf{y}_0}{2}$ , we know  $F(0, \cdot)$  explicitly. Even in the generality of the hypotheses of Lemma 19, one can obtain the following necessary condition on the drift of a Euclidean diffusion for existence of a Markovian maximal coupling: use (ii) of the above lemma and take  $t \downarrow 0$ .

**Corollary 21** *Under the hypotheses of Lemma 19 and (34),  $\mathbf{n}(0)$  must be an eigenvector of  $\frac{S(\mathbf{x}) + S(F(0, \mathbf{x}))}{2}$  corresponding to some eigenvalue  $\lambda(\mathbf{x})$ , for every  $\mathbf{x} \in \mathbb{R}^d$ .*

Briefly restrict attention to the case where  $\mathbf{b}(\mathbf{x})$  is affine in  $\mathbf{x}$ . The following theorem completely classifies the set of such drifts which ensure Markovian maximal coupling.

**Theorem 22** *Assume (A1), (A2). Let  $\mathbf{b}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$  for some  $(d \times d)$  matrix  $A$  and some  $d$ -dimensional vector  $\mathbf{c}$ . Denote  $S = \frac{A + A^\top}{2}$  and  $T = \frac{A - A^\top}{2}$ . Then a Markovian maximal coupling  $(X, Y)$  exists from starting points  $\mathbf{x}_0$  and  $\mathbf{y}_0$  if and only if there exists an eigenvalue  $\lambda_0$  of  $S$  such that the vectors  $T^k(\mathbf{x}_0 - \mathbf{y}_0)$  (for  $0 \leq k \leq d - 1$ ) all lie in the eigenspace of  $S$  corresponding to  $\lambda_0$ . In this case (using matrix exponentials  $\exp$ ),*

$$\mathbf{n}(t) = \exp(Tt) \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}, \text{ and} \quad (42)$$

$$l(t) = e^{\lambda_0 t} \frac{|\mathbf{x}_0|^2 - |\mathbf{y}_0|^2}{2|\mathbf{x}_0 - \mathbf{y}_0|} + e^{\lambda_0 t} \int_0^t \frac{(\mathbf{x}_0 - \mathbf{y}_0)^\top}{|\mathbf{x}_0 - \mathbf{y}_0|} \exp(-(T + \lambda_0 \mathbb{I})s) \mathbf{c} \, ds. \quad (43)$$

*Proof* Suppose there exists a Markovian maximal coupling  $(X, Y)$  starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . From (ii) and (iii) of Lemma 20 we get the following:

$$S\mathbf{n}(t) = \lambda(t)\mathbf{n}(t) \quad (44)$$

(where we note that  $\lambda$  is a function of  $t$  only) and

$$T\mathbf{n}(t) = \dot{\mathbf{n}}(t). \quad (45)$$

Solving (45), we get (42). Since  $T$  is skew-symmetric, the above formula implies  $|\mathbf{n}(t)| = 1$  for all  $t$ .

The finite symmetric matrix  $S$  has discrete spectrum; by this, and the continuity of  $\mathbf{n}(\cdot)$  and  $\lambda(\cdot)$ , it follows immediately from (44) that  $\lambda(\cdot) \equiv \lambda_0$  for some constant  $\lambda_0$ . Thus  $\mathbf{n}(t)$ , as given by (42), must lie in the eigenspace of  $S$  corresponding to  $\lambda_0$ , for all time  $t$ . Substituting this formula for  $\mathbf{n}(t)$  in Eq. (44) and differentiating (42)  $k$  times with respect to  $t$  (for  $k = 0, 1, \dots, d - 1$ ), then setting  $t = 0$ , we obtain that the vectors  $T^k(\mathbf{x}_0 - \mathbf{y}_0)$  for  $0 \leq k \leq d - 1$  must all lie in the eigenspace of  $S$  corresponding to  $\lambda_0$ . As  $T$  solves its characteristic equation, it is clear that all the higher powers  $T^k(\mathbf{x}_0 - \mathbf{y}_0)$  for  $k \geq d$  must also lie in this eigenspace. Using the series representation of  $\exp(Tt)$ , this means that  $\mathbf{n}(t)$  must also lie in this eigenspace for all  $t$ .

To solve for  $l$ , note that computation with (33), (34), (41) yields the following expression for  $\mathbf{n} = \mathbf{n}(t)$  and  $l = l(t)$ :

$$2(\dot{l}\mathbf{n} - l\dot{\mathbf{n}}) + 2l(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)A\mathbf{n} - 2\mathbf{n}\mathbf{n}^\top\mathbf{c} = 0. \quad (46)$$

On the other hand, (44) and (45) yield

$$A\mathbf{n} = \lambda_0\mathbf{n} + \dot{\mathbf{n}}.$$

Substituting into (46) and simplifying,

$$\dot{l} = \lambda_0 l + \mathbf{n}^\top \mathbf{c}. \quad (47)$$

Solving this equation, using the solution for  $\mathbf{n} = \mathbf{n}(t)$  obtained from (42), we get (43).

Conversely, suppose there exists an eigenvalue  $\lambda_0$  of  $S$  such that the vectors  $T^k(\mathbf{x}_0 - \mathbf{y}_0)$  (for  $0 \leq k \leq d-1$ ) all lie in the eigenspace of  $S$  corresponding to  $\lambda_0$ . To prove the existence of a Markovian maximal coupling  $(X, Y)$  starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$ , we will show that (34) holds with  $\mathbf{n}$  and  $l$  as given in the theorem.

Clearly, for this choice of  $\mathbf{n}$  and  $l$ , (45) and (47) hold. Using these, we obtain

$$\dot{\mathbf{n}}\mathbf{n}^\top - \mathbf{n}\dot{\mathbf{n}}^\top = T\mathbf{n}\mathbf{n}^\top + \mathbf{n}\mathbf{n}^\top T$$

and

$$\dot{l}\mathbf{n} - l\dot{\mathbf{n}} = \lambda_0 l\mathbf{n} + \mathbf{n}\mathbf{n}^\top \mathbf{c} - lT\mathbf{n}.$$

Now, observe that  $S\mathbf{n} = \lambda_0\mathbf{n}$  and

$$\mathbf{n}^\top A\mathbf{n} = \mathbf{n}^\top S\mathbf{n} = \lambda_0.$$

Using these, we can write

$$\begin{aligned} (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{b}(F(t, \mathbf{x})) &= (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)(A(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{x} + 2lA\mathbf{n} + \mathbf{c}) \\ &= (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)A(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{x} - 2\lambda_0 l\mathbf{n} + 2lT\mathbf{n} \\ &\quad + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)\mathbf{c}. \end{aligned}$$

Applying the above relations, the right hand side of (34) becomes

$$\begin{aligned} &[2(T\mathbf{n}\mathbf{n}^\top + \mathbf{n}\mathbf{n}^\top T) + (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)A(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)]\mathbf{x} + \mathbf{c} \\ &= (A\mathbf{x} + \mathbf{c}) + [-2S\mathbf{n}\mathbf{n}^\top - 2\mathbf{n}\mathbf{n}^\top S + 4(\mathbf{n}^\top A\mathbf{n})\mathbf{n}\mathbf{n}^\top]\mathbf{x}, \end{aligned}$$

where we used  $A = S + T$ . Now, using  $S\mathbf{n} = \lambda_0\mathbf{n}$  and  $\mathbf{n}^\top A\mathbf{n} = \lambda_0$  again, we get

$$-2S\mathbf{n}\mathbf{n}^\top - 2\mathbf{n}\mathbf{n}^\top S + 4(\mathbf{n}^\top A\mathbf{n})\mathbf{n}\mathbf{n}^\top = 0,$$

and thus, (34) holds, proving the theorem.  $\square$

The following corollary is immediate from the above theorem.

**Corollary 23** *If  $d = 2$ , then under the hypotheses of Theorem 22,  $A$  is either a symmetric matrix or of the form  $\lambda_0 \mathbb{I} + T$  for some real scalar  $\lambda_0$  and a skew-symmetric matrix  $T$ .*

*Proof* If the skew-symmetric part  $T$  of  $A$  is non-zero, then  $\mathbf{x}_0 - \mathbf{y}_0$  and  $T(\mathbf{x}_0 - \mathbf{y}_0)$  are non-zero, mutually orthogonal vectors which lie in the eigenspace of  $S$  corresponding to  $\lambda_0$ . Thus, this eigenspace is the whole of  $\mathbb{R}^2$  and  $S = \lambda_0 \mathbb{I}$ .  $\square$

Now, we state and prove the main theorem of this section. Recall the Local Perturbation condition *LPC* described in the introduction.

**Theorem 24** *Assume (A1) and (A2) hold for a time-homogeneous Euclidean diffusion. Then LPC holds if and only if there exist a real scalar  $\lambda_0$ , a skew-symmetric matrix  $T$  and a vector  $\mathbf{c} \in \mathbb{R}^d$  such that the diffusion drift is given by*

$$\mathbf{b}(\mathbf{x}) = \lambda_0 \mathbf{x} + T\mathbf{x} + \mathbf{c}$$

for all  $\mathbf{x} \in \mathbb{R}^d$ .

*Proof* We need to show that the set of eigenvalues of  $S(\mathbf{x})$  for any  $\mathbf{x} \in \mathbb{R}^d$  is the singleton  $\{\lambda_0\}$  and the skew-symmetric part  $T(\mathbf{x})$  is a constant matrix  $T$ . Write

$$\mathcal{H}_0 = \{H(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r), \mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)\}.$$

Our approach is to choose an appropriate set of mirrors  $\mathcal{H} \subseteq \mathcal{H}_0$  and then to consider the orbit of a point  $\mathbf{z} \in \mathbb{R}^d$  under repeated reflections in this set of mirrors, defined as

$$\mathcal{O}(\mathbf{z}) = \left\{ \mathbf{w} \in \mathbb{R}^d : \text{there exist } h_1, \dots, h_k \in \mathcal{H} \text{ such that } \mathbf{w} = h_k \dots h_1 \mathbf{z} \right\}.$$

We then use the constraint relations between a point and its reflection obtained in Lemma 20.

This idea is made more precise in the following internal lemmas.

**Lemma 25** *Under the hypotheses of Theorem 24, there exists  $\lambda_0 \in \mathbb{R}$  such that  $S(\mathbf{x}) = \lambda_0 \mathbb{I}$  for all  $\mathbf{x} \in \mathbb{R}^d$ .*

*Proof* Suppose that  $X$  and  $Y$  start at  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  respectively. It follows from letting  $t \downarrow 0$  in part (i) of Lemma 20 that, for all  $\mathbf{z} \in \mathbb{R}^d$ ,  $S(\mathbf{z})$  and  $S(H(\mathbf{x}, \mathbf{y})\mathbf{z})$  have the same set of eigenvalues. (Recall that  $H(\mathbf{x}, \mathbf{y})\mathbf{z}$  represents reflection of  $\mathbf{z}$  in the hyperplane  $H(\mathbf{x}, \mathbf{y})$ .)

Denote  $\mathbf{x}^* = (\mathbf{x}_0 + \mathbf{y}_0)/2$  and let  $\mathbf{v}_1 = \mathbf{x}_0 - \mathbf{x}^*$ . Extend  $\mathbf{v}_1$  to a basis  $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ . If  $\varepsilon$  is sufficiently small then the linearly independent vectors  $\mathbf{n}_i = \mathbf{v}_1 + \varepsilon \mathbf{v}_i$ ,  $i = 1, \dots, d$  are such that  $\{\mathbf{x}^* + \mathbf{n}_i : i = 1, \dots, d\} \subset \mathcal{B}(\mathbf{x}_0, r)$  and  $\{\mathbf{x}^* - \mathbf{n}_i : i = 1, \dots, d\} \subset \mathcal{B}(\mathbf{y}_0, r)$ . Defining  $\mathbf{x}_i = \mathbf{x}^* + \mathbf{n}_i$  and  $\mathbf{y}_i = \mathbf{x}^* - \mathbf{n}_i$ , it follows that  $\mathbf{x}^* \in H(\mathbf{x}_i, \mathbf{y}_i)$  for all  $i$ . For each  $i$ , consider maximally coupled diffusions begun at  $(\mathbf{x}_i, \mathbf{y}_i)$ : applying part (ii) of Lemma 20 and letting  $t \downarrow 0$ , it follows that  $\mathbf{n}_i$  is an eigenvector of  $S(\mathbf{x}^*)$ . By construction, no  $\mathbf{n}_i$  is orthogonal to any other  $\mathbf{n}_j$ . Since  $S(\mathbf{x}^*)$  is symmetric, it



follows that  $\{\mathbf{n}_i : i = 1, \dots, d\}$  correspond to the same eigenvalue, say  $\lambda_0$  and thus,  $S(\mathbf{x}^*) = \lambda_0 \mathbb{I}$ .

Choosing the set of mirrors  $\mathcal{H} = \mathcal{H}_0$ , consider the orbit  $\mathcal{O}(\mathbf{x}^*)$  of  $\mathbf{x}^*$  in  $\mathcal{H}$ . If  $\mathcal{O}(\mathbf{x}^*) = \mathbb{R}^d$ , then the lemma follows from the previous observation that for any  $\mathbf{z} \in \mathcal{O}(\mathbf{x}^*)$ , the set of eigenvalues of  $S(\mathbf{z})$  agrees with that of  $S(\mathbf{x}^*)$ .

To see this, let  $L$  be the line that passes through  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . Let  $\mathbf{v}_0 = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$ . Write  $\mathbf{x}_\delta = \mathbf{x}_0 + \delta \mathbf{v}_0$  and  $\mathbf{y}_\delta = \mathbf{y}_0 + \delta \mathbf{v}_0$  for all  $\delta \in (-r, r)$ . Thus the mirrors  $h_\delta = H(\mathbf{x}_\delta, \mathbf{y}_\delta) \in \mathcal{H}$  for all such  $\delta$ , and the orbit of  $\mathbf{x}^*$  under reflection in  $\{h_\delta : \delta \in (-r, r)\}$  is the whole of  $L$ . Thus  $L \subseteq \mathcal{O}(\mathbf{x}^*)$ .

Now, for any  $\mathbf{z} \in \mathbb{R}^d$ , let  $H$  be a plane (dimension of  $H$  is two) containing the line  $L$  and the point  $\mathbf{z}$ . For sufficiently small  $\varepsilon > 0$ , for all  $\delta \in (-\varepsilon, \varepsilon)$  the mirror  $h'_\delta$  containing  $\mathbf{x}^*$  and having normal vector  $\mathbf{v}_\delta \in H$  and making an angle  $\delta$  with  $\mathbf{v}_0$  lies in  $\mathcal{H}$ . Denote by  $C$  the circle centred at  $\mathbf{x}^*$ , lying in  $H$  and passing through  $\mathbf{z}$ . Let  $\boldsymbol{\theta} \in L \cap C$ . Then the orbit of  $\boldsymbol{\theta}$  under reflection in  $\{h'_\delta : \delta \in (-\varepsilon, \varepsilon)\}$  is the whole of  $C$ . In particular,  $\mathbf{z} \in \mathcal{O}(\mathbf{x}^*)$ . This shows that  $\mathcal{O}(\mathbf{x}^*) = \mathbb{R}^d$  and the lemma follows.  $\square$

Before proceeding further with the proof of Theorem 24, we record a general fact about real skew-symmetric matrices which follows by spectral decomposition [13].

**Lemma 26** *If  $\mathcal{N}$  is the null space of a  $(d \times d)$  real skew-symmetric matrix  $T$ , then  $d - \dim(\mathcal{N})$  is even.*

We now show that the skew-symmetric part  $T(\mathbf{x})$  is a constant matrix  $T$ .

**Lemma 27** *Under the hypotheses of Theorem 24,  $T(\mathbf{x}) \equiv T$  for all  $\mathbf{x} \in \mathbb{R}^d$ .*

*Proof* The proof breaks into three steps.

**Step 1.** *If  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$ , then for all  $\mathbf{z}, \mathbf{z}' \in H(\mathbf{x}, \mathbf{y})$ ,  $T(\mathbf{z}) = T(\mathbf{z}')$ . Set  $\mathbf{z}^* = \frac{\mathbf{z} + \mathbf{z}'}{2}$ ,  $\mathbf{v}_1 = \frac{\mathbf{z} - \mathbf{z}'}{|\mathbf{z} - \mathbf{z}'|}$  and  $\mathbf{v}_2 = \frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|}$ . Extend  $\mathbf{v}_1, \mathbf{v}_2$  to an orthonormal basis  $\mathbf{v}_1, \dots, \mathbf{v}_d$  of  $\mathbb{R}^d$ . Using the method of the proof of Lemma 25, construct independent vectors  $\mathbf{n}_i = \mathbf{v}_2 + \varepsilon \mathbf{v}_i$ ,  $i = 2, \dots, d$ , choosing  $\varepsilon > 0$  small enough so that*

$$H(\mathbf{z}^* + \mathbf{n}_i, \mathbf{z}^* - \mathbf{n}_i) \mathbf{x} \in \mathcal{B}(\mathbf{y}_0, r)$$

for all  $i = 2, \dots, d$ . Writing  $\mathbf{x}_i = \mathbf{z}^* + \mathbf{n}_i$  and  $\mathbf{y}_i = \mathbf{z}^* - \mathbf{n}_i$ , and with a possibly smaller choice of  $\varepsilon > 0$ , the hyperplane  $H(\mathbf{x}_i, \mathbf{y}_i)$  lies in  $\mathcal{H}_0$  and the line joining  $\mathbf{z}$  and  $\mathbf{z}'$  is contained in  $H(\mathbf{x}_i, \mathbf{y}_i)$  for all  $i = 2, \dots, d$ . Thus,  $H(\mathbf{x}_i, \mathbf{y}_i)\mathbf{z} = \mathbf{z}$  and  $H(\mathbf{x}_i, \mathbf{y}_i)\mathbf{z}' = \mathbf{z}'$  for all  $i = 2, \dots, d$ . Taking  $t \downarrow 0$  in part (iii) of Lemma 20, it follows that

$$(T(\mathbf{z}) - T(\mathbf{z}'))\mathbf{n}_i = 0$$

for all  $i = 2, \dots, d$ , implying  $d - \mathcal{N}(T(\mathbf{z}) - T(\mathbf{z}')) \leq 1$ . Together with Lemma 26, this establishes Step 1.

**Step 2.** There is  $\varepsilon > 0$  such that  $T(\mathbf{z}) = T(\mathbf{z}')$  for all  $\mathbf{z}, \mathbf{z}' \in \{\mathbf{w} \in \mathbb{R}^d : \text{dist}(\mathbf{w}, H(\mathbf{x}_0, \mathbf{y}_0)) < \varepsilon\}$ , where  $\text{dist}(\mathbf{w}, A)$  denotes the distance of  $\mathbf{w}$  from the set  $A$ .

Choose  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  such that the vector  $\mathbf{x} - \mathbf{y}_0$  is not parallel to  $\mathbf{x}_0 - \mathbf{y}_0$ . It follows from Step 1 that  $T(\mathbf{z}) = T(\mathbf{z}')$  for all  $\mathbf{z}, \mathbf{z}' \in H(\mathbf{x}, \mathbf{y})$ . Choose  $\varepsilon > 0$  such that  $\mathbf{y}_\delta = \mathbf{y}_0 + \delta \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|} \in \mathcal{B}(\mathbf{y}_0, r)$  for all  $\delta \in (-2\varepsilon, 2\varepsilon)$ . Note that the vector  $\mathbf{x} - \mathbf{y}_\delta$  is not parallel to  $\mathbf{x}_0 - \mathbf{y}_0$  for any  $\delta \in (-2\varepsilon, 2\varepsilon)$ . Using Step 1 again,  $T(\mathbf{z}) = T(\mathbf{z}')$  for all  $\mathbf{z}, \mathbf{z}' \in H(\mathbf{x}_0, \mathbf{y}_\delta)$ . The assertion now follows from Step 1 and the fact that  $H(\mathbf{x}_0, \mathbf{y}_\delta) \cap H(\mathbf{x}, \mathbf{y})$  is non-empty for each  $\delta \in (-2\varepsilon, 2\varepsilon)$ .

**Step 3.** Now we work with the set of mirrors

$$\mathcal{H} = \{H(\mathbf{x}_0, \mathbf{y}_\delta) : \delta \in (-2\varepsilon, 2\varepsilon)\},$$

where  $\varepsilon$  is chosen as in Step 2. For notational convenience, we write  $h_\delta = H(\mathbf{x}_0, \mathbf{y}_\delta)$ . The  $\mathbf{y}_\delta = \mathbf{y}_0 + \delta \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  all lie on the same line through  $\mathbf{x}_0$ , and therefore all these mirrors have a common normal vector, which we write  $\mathbf{n}^*$ . Let  $(l_\delta, \mathbf{n}_\delta)$  parametrize the interface  $I(\mathbf{x}_0, \mathbf{y}_\delta, \cdot)$  corresponding to the starting points  $\mathbf{x}_0$  and  $\mathbf{y}_\delta$  of the diffusions  $X$  and  $Y$  respectively. For each  $\delta$ ,  $\mathbf{n}_\delta(0) = \mathbf{n}^*$ . Furthermore, by letting  $t \downarrow 0$  in part (iii) of Lemma 20,

$$\dot{\mathbf{n}}_\delta(0) = T\left(\frac{\mathbf{x}_0 + \mathbf{y}_\delta}{2}\right) \mathbf{n}^*.$$

Given  $\delta \in (-2\varepsilon, 2\varepsilon)$ , the distance of the point  $\frac{\mathbf{x}_0 + \mathbf{y}_\delta}{2}$  from the hyperplane  $H(\mathbf{x}_0, \mathbf{y}_0)$  is less than  $\varepsilon$ . Consequently Step 2 implies that  $\dot{\mathbf{n}}_\delta(0) = \dot{\mathbf{n}}_0(0) = \mathbf{n}'$  (say) for all  $\delta \in (-2\varepsilon, 2\varepsilon)$ .

Choose any  $\mathbf{z}, \mathbf{z}' \in \mathbb{R}^d$  such that  $\mathbf{z}' = \mathbf{z} + \delta \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  for some  $\delta \in (-2\varepsilon, 2\varepsilon)$ . Set  $\mathbf{z}^* = h_0 \mathbf{z}$  so that  $\mathbf{z} = h_0 \mathbf{z}^*$ . Noting that  $\mathbf{z}, \mathbf{z}^*, \mathbf{z}'$  lie on the same line perpendicular to  $H(\mathbf{x}_0, \mathbf{y}_0)$ , it follows from an argument about one-dimensional reflections that  $\mathbf{z}' = h_\delta \mathbf{z}^*$ .

Then, by part (i) of Lemma 20, we get

$$\begin{aligned} T(\mathbf{z}^*) &= 2(\mathbf{n}' \mathbf{n}^{*\top} - \mathbf{n}^* \mathbf{n}'^\top) + (\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top})T(\mathbf{z})(\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top}) \\ &= 2(\mathbf{n}' \mathbf{n}^{*\top} - \mathbf{n}^* \mathbf{n}'^\top) + (\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top})T(\mathbf{z}')(\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top}) \end{aligned} \quad (48)$$

from which we get

$$(\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top})(T(\mathbf{z}) - T(\mathbf{z}'))(\mathbb{I} - 2\mathbf{n}^* \mathbf{n}^{*\top}) = 0$$

which gives  $T(\mathbf{z}) = T(\mathbf{z}')$ . Hence the lemma follows.  $\square$

Lemmas 25 and 27 together are sufficient to prove Theorem 24.  $\square$

Theorem 24 can be strengthened if  $\dot{\mathbf{n}}(t) = 0$  for all  $t$ , i.e., the interface *translates but does not rotate in time*. We state this in the following theorem. Since there is no rotation, the driving Brownian motions in the stochastic differential equation for  $X$  and

$Y$  are constant reflections of each other. So we can assume without loss of generality that  $l(0) = 0$  and  $\mathbf{n}(t) \equiv \mathbf{e}_1$ .

**Theorem 28** Assume (A1) and (A2) hold for a time-homogeneous Euclidean diffusion. Suppose there exists a Markovian maximal coupling of  $X$  and  $Y$  starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$  respectively, such that the interface  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  is parametrized by  $((l(t), \mathbf{e}_1) : t \geq 0)$  with  $l(0) = 0$ . Then there are only two possibilities:

- (i)  $l(t) = 0$  for all  $t \geq 0$ , in which case the drift vectorfield  $\mathbf{b}$  must satisfy

$$\mathbf{b}(h_1 \mathbf{x}) = h_1 \mathbf{b}(\mathbf{x})$$

for all  $\mathbf{x} \in \mathbb{R}^d$ .

- (ii)  $l(t) \neq 0$  for some  $t > 0$ , in which case the drift vectorfield  $\mathbf{b}$  must satisfy

$$\mathbf{b}(x_1, \mathbf{x}^{(1)}) = \left( c_1 x_1 + c_2, \mathbf{f}(\mathbf{x}^{(1)}) \right)^\top$$

for all  $\mathbf{x} = (x_1, \mathbf{x}^{(1)}) \in \mathbb{R}^d$ , where  $c_1, c_2$  are constants and  $\mathbf{f} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^{d-1}$  is continuously differentiable.

*Proof* Part (i) follows from the fact that the generators of  $Y$  and  $h_1 X$  are the same.

To prove part (ii), note that by part (i) of Lemma 20:

$$\nabla \mathbf{b}(x_1, \mathbf{x}^{(1)}) = \begin{bmatrix} \partial_1 b_1(x_1, \mathbf{x}^{(1)}) & \mathbf{0} \\ \mathbf{0} & \nabla^{(1)} \mathbf{b}^{(1)}(x_1, \mathbf{x}^{(1)}) \end{bmatrix}. \quad (49)$$

for all  $\mathbf{x} = (x_1, \mathbf{x}^{(1)}) \in \mathbb{R}^d$ , where  $\mathbf{b}^{(1)} = (b_2, \dots, b_d)^\top$  and  $\nabla^{(1)}$  denotes partial derivatives with respect to the variables of  $\mathbf{x}^{(1)}$ . From (49), we deduce that  $b_1(x_1, \mathbf{x}^{(1)}) = f_1(x_1)$  and  $\mathbf{b}^{(1)}(x_1, \mathbf{x}^{(1)}) = \mathbf{f}(\mathbf{x}^{(1)})$  for continuously differentiable functions  $f_1 : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  and  $\mathbf{f} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^{d-1}$ .

We may assume that (without loss of generality)  $(0, \varepsilon) \subset \text{Range}(l)$  for some  $\varepsilon > 0$ . Choose the set of mirrors

$$\mathcal{H} = \{H(\mathbf{x}_0, \mathbf{y}_\delta) : \delta \in (0, \varepsilon)\}$$

where, as before,  $\mathbf{y}_\delta = \mathbf{x}_0 + \delta \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$ . Now, iterated reflections in  $\mathcal{H}$  as in the proof of Theorem 24 yield  $f'_1(x_1 + a) = f'_1(x_1)$  for all  $x_1, a \in \mathbb{R}$ . Hence,  $f'_1(x_1) = c_1$  for all  $x_1 \in \mathbb{R}$ , for some constant  $c_1$ . Thus,  $\mathbf{b}$  has to be of the required form.  $\square$

The case of one-dimensional diffusions is a trivial consequence of the above theorem, as noted in the next corollary.

**Corollary 29** Assume (A1) and (A2) hold for a one-dimensional time-homogeneous Euclidean diffusion. Then there exists a Markovian maximal coupling of  $X$  and  $Y$  starting from  $x_0$  and  $y_0$  respectively if and only if either the drift vectorfield  $b$  is affine or it obeys the reflection symmetry  $b(x) = -b(x_0 + y_0 - x)$  for all  $x \in \mathbb{R}$ .

**Remark 30** Corollary 29 completely characterises all one-dimensional time-homogeneous diffusions subject to the regularity conditions (A1) and (A2) and permitting Markovian maximal couplings, even with a varying twice-continuously-differentiable diffusion coefficient  $\sigma(\cdot) : \mathbb{R} \rightarrow [c, \infty)$  for some  $c > 0$ . Let  $X$  be given by

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t \quad (50)$$

and similarly for  $Y$ . Define the function

$$F(x) = \int_0^x \frac{1}{\sigma(z)} dz,$$

and set  $U_t = F(X_t)$ . Then, it follows from Itô calculus that

$$dU_t = dB_t + \left( \frac{b \circ F^{-1}(U_t)}{\sigma \circ F^{-1}(U_t)} - \frac{\sigma' \circ F^{-1}(U_t)}{2} \right) dt. \quad (51)$$

Thus, the conditions on  $b$  derived in the case  $\sigma \equiv 1$  readily carry over to conditions on the drift term of (51) for general  $\sigma$ .

### 3 Markovian maximal couplings for manifolds

In this section, we analyse rigidity phenomena for Markovian maximal couplings (MMC) for smooth elliptic diffusions, and demonstrate that there are powerful geometric consequences arising from a natural connection to the theory of diffusion processes on manifolds (specifically, the notion of Riemannian Brownian motion with drift). The main task of this section is to understand how the Euclidean arguments of Sect. 2 carry over to the manifold case. In particular, the existence of Markovian maximal couplings (together with LPC) has profound rigidity consequences for the geometry of the manifold.

We commence by summarizing the Riemannian geometry required to establish these consequences. Let  $M$  be a connected smooth manifold of dimension  $d$  (the results which follow are actually significant even in the case when  $M = \mathbb{R}^d$ ). Following [11], a strong Markov process  $X$  on  $M$  is said to be a diffusion process if each  $C^2$  function  $f$  belongs to the domain of definition of the *characteristic operator*  $L$  given by

$$Lf(\mathbf{x}) = \lim_{N \downarrow \mathbf{x}} \frac{\mathbb{E}_{\mathbf{x}}[f(X_{\tau_N})] - f(\mathbf{x})}{\mathbb{E}_{\mathbf{x}}[\tau_N]} \quad (52)$$

where  $N$  denotes a system of neighbourhoods shrinking to  $\mathbf{x}$ ,  $\tau_N$  denotes the first exit time from  $N$  and  $\mathbb{E}$  denotes expectation with respect to the measure induced by the Markov process. In any local system of coordinates  $(x^1, \dots, x^d)$ , the operator  $L$  takes the form

$$Lf(\mathbf{x}) = \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial^2 f}{\partial x^i \partial x^j} + \sum_{i=1}^d v_i(\mathbf{x}) \frac{\partial f}{\partial x^i} \quad (53)$$

where the *diffusion matrix*  $A = \{a_{ij}\}$  is non-negative definite and  $\{v_i\}$  denotes the *drift vectorfield*. We will assume  $a_{ij}$  and  $v_i$  are smooth functions. Note that the general form of the operator does not depend on the specific choice of coordinates. We call  $X$  an *elliptic diffusion* if  $L$  is an elliptic operator (in other words, if  $A$  is positive-definite). As in the previous section, we deal only with elliptic diffusions.

Following [29], if we furnish  $M$  with the Riemannian metric  $g$  which is given in local coordinates by  $g_{ij} = (A^{-1})_{ij}$  then the operator  $L$  can be rewritten in the form

$$L = \frac{1}{2} \Delta_M + \mathbf{b} \quad (54)$$

where  $\Delta_M$  is the *Laplace–Beltrami operator* for the Riemannian metric, and  $\mathbf{b}$  is the (intrinsic) drift vectorfield. When  $\mathbf{b} = 0$ , the corresponding Markov process is called *Brownian motion* on  $M$ . Thus, we see that any diffusion process on  $M$  can be written as ‘Brownian motion plus drift’ if  $M$  is given a suitable metric. Henceforth, we will assume that  $M$  is endowed with this metric  $g$ , so that we can view  $M$  as a smooth Riemannian manifold  $(M, g)$ .

*Note* Throughout this section, we will make the following assumptions:

- (i) The Riemannian manifold  $(M, g)$  obtained above is *complete* (we say that the diffusion  $X$  is *diffusion-geodesic complete*). This is a purely technical assumption and the completeness is usually not too hard to check as we know the diffusion coefficients explicitly. In particular, diffusion-geodesic completeness trivially holds on compact manifolds. Diffusion-geodesic completeness is not a necessary condition for the existence of Markovian maximal couplings, as can be seen for dimension  $d \geq 2$  by considering reflection couplings of Brownian motions on the  $d$ -dimensional punctured sphere  $\mathbb{S}^d - \{P\}$  obtained by deleting a point  $P$  from the sphere  $\mathbb{S}^d$  (and the corresponding couplings of diffusions obtained on the plane by stereographic projection). In this example, the existence of a rich supply of MMC follows from the fact that this space has a completion  $\mathbb{S}^d$  on which we can construct MMC of Brownian motions started from any two points (see [25]), and from the fact that if  $d \geq 2$  then the Brownian motion started in  $\mathbb{S}^d - \{P\}$  almost surely does not hit  $P$ . It is an interesting question whether this is the ‘generic’ example for instances where diffusion-geodesic completeness fails but Markovian maximal couplings exist, raising issues which seem somewhat reminiscent of the topic of resolution of singularities in algebraic geometry. We hope to address this in a future article.
- (ii) Our diffusion process  $X$  is *defined for all time*. This is to ensure that we are dealing with probability densities which is essential for the arguments in Sect. 1.1 to go through. For Brownian motion on  $M$ , this can be resolved by ensuring that  $M$  is *stochastically complete*. There are a number of intrinsic geometric properties of  $M$  that ensure stochastic completeness, such as the existence of a constant lower bound on the Ricci curvature. See [17], for example, for more details.

Let  $\mathcal{G} = \text{Iso}(M)$  denote the *group of (global) isometries* of  $M$ . This can be shown to be a Lie group [30], and it plays an important rôle in the following arguments. As  $M$  is complete and connected, any pair of points in  $M$  are connected by a geodesic. Furthermore, there are no branching geodesics in Riemannian manifolds. (More details on these geometric notions can be found in [5, 7].)

### 3.1 Brownian motion with drift on the manifold

Not only can any smooth elliptic diffusion on  $M$  be written as Brownian motion with drift on  $(M, g)$ , but also this permits a rather explicit geometric construction of the diffusion which facilitates the discussion of probabilistic coupling techniques, namely the *Eells–Elworthy–Malliavin* construction [12].

Using terminology expounded (for example) in [17], let  $\mathcal{O}_x(M)$  denote the set of orthonormal frames of the tangent space  $T_x M$ . The *orthonormal frame bundle*

$$\mathcal{O}(M) = \bigcup_{x \in M} \mathcal{O}_x(M)$$

possesses a natural smooth manifold structure of dimension  $\frac{d(d+1)}{2}$ . Denote the canonical projection map by  $\pi : \mathcal{O}(M) \rightarrow M$ .

A curve  $u$  in  $\mathcal{O}(M)$  is said to be *horizontal* if  $u_t$  is the parallel transport (associated with the *Levi–Civita* connection) of the frame  $u_0$  along the curve  $\pi u_t$ . For each  $u \in \mathcal{O}(M)$ , the tangent space  $T_u \mathcal{O}(M)$  can be expressed as a direct sum

$$T_u \mathcal{O}(M) = V_u \mathcal{O}(M) \oplus H_u \mathcal{O}(M),$$

where  $V_u \mathcal{O}(M)$  is a  $\frac{d(d-1)}{2}$ -dimensional vector space corresponding to the isotropy group (frame rotations) at  $\pi u$ , and the  $d$ -dimensional vector space  $H_u \mathcal{O}(M)$  is the space of tangent vectors of horizontal curves passing through  $u$ .

For each  $u \in \mathcal{O}(M)$ , let  $H_i(u)$  denote the unique horizontal vector lying in  $H_u \mathcal{O}(M)$  such that

$$\pi_* H_i(u) = u e_i,$$

where  $u e_i$  denotes the  $i$ th unit vector of the orthonormal frame  $u$ .

This framework provides an expressive way to define smooth elliptic diffusions (and other semimartingale processes) on  $M$ , as follows.

Let  $\mathbf{b}$  be a smooth vectorfield on  $M$ . This yields a natural vectorfield  $\mathbf{B}$  on  $\mathcal{O}(M)$  given by

$$\mathbf{B}(u) = \sum_i b_i(u) H_i(u), \quad (55)$$

where  $b_i(u) = \langle \mathbf{b}(\pi u), u e_i \rangle_{\pi u}$  (here  $\langle \cdot, \cdot \rangle$  denotes the Riemannian inner product). We will call this the *lifted drift*. Consider the following Stratonovich differential equation on  $\mathcal{O}(M)$ :

$$dU_t = \sum_i H_i(U_t) \circ dW_t^i + \mathbf{B}(U_t) dt. \quad (56)$$

where  $W$  is a  $d$ -dimensional Euclidean Brownian motion. The diffusion on  $M$  with drift  $\mathbf{b}$  is obtained simply as the projection  $X_t = \pi U_t$ . The pivotal fact justifying this construction is that we can define a second order operator on  $\mathcal{O}(M)$  (Bochner's horizontal Laplacian) given by

$$\Delta_{\mathcal{O}(M)} = \sum_{i=1}^d H_i^2$$

such that the Laplace-Beltrami operator  $\Delta_M$  on  $M$  satisfies

$$\Delta_M f(\mathbf{x}) = \Delta_{\mathcal{O}(M)} f \circ \pi(u)$$

for any  $u \in \mathcal{O}(M)$  such that  $\pi u = \mathbf{x}$ . The generator  $L$  of the diffusion  $X$  defined at the start of Sect. 3 satisfies

$$Lf(\mathbf{x}) = \frac{1}{2} \Delta_M f(\mathbf{x}) + \mathbf{b}f(\mathbf{x}) \quad (57)$$

for any  $u \in \mathcal{O}(M)$  such that  $\pi u = \mathbf{x}$ , and any  $C^2$  test function  $f$  on  $M$ .

Note that, when  $\mathbf{b} = 0$ , the above construction reduces to the classical Eells-Elworthy-Malliavin construction of Brownian motion on  $M$ .

### 3.2 Couplings of diffusions on manifolds

Once we have the above construction, a natural question to ask is: when is there a Markovian maximal coupling (MMC) for two copies of the diffusion starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$ ? In the Euclidean case there is a complete characterization of the class of time-homogeneous diffusions under *LPC*, which is to say, when two copies of the diffusion can be maximally coupled whenever they start from  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  (for  $\mathcal{B}(\mathbf{x}_0, r)$  and  $\mathcal{B}(\mathbf{y}_0, r)$  chosen to be two arbitrary disjoint open balls in  $\mathbb{R}^d$ ). Theorem 24 shows that the class of such diffusions is actually very small.

The proof of Theorem 24 depends strongly on a wealth of isometries of Euclidean space arising *via* iterated reflections. Very few other  $d$ -dimensional Riemannian manifolds have many isometries, and so we may expect an even stronger rigidity phenomenon to hold for the geometry of (non-Euclidean) manifolds on which there is a good supply of MMC. The work of this section substantiates this expectation.

We begin by recalling briefly some notions from the Euclidean case (Sect. 2). We have noted that the Local Perturbation Condition *LPC* (Definition 1) makes sense for any metric space, including the Riemannian manifold case. Let  $X$  and  $Y$  be two copies of the elliptic diffusion derived from the stochastic differential equation (56), and starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$  respectively. Note that the assumptions of ellipticity and smoothness of the coefficients of  $L$  together ensure that the law of  $X$  (equivalently  $Y$ )

has a smooth positive density with respect to the Riemannian volume measure  $m$  for every positive time  $t > 0$ , which we write as  $p(\mathbf{x}_0; t, \mathbf{z})$ ,  $p(\mathbf{y}_0; t, \mathbf{z})$  for  $t > 0$ ,  $\mathbf{z} \in M$ .

We suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold for the regular elliptic diffusion  $X$ , so that the resulting Riemannian manifold  $M$  is geodesically complete and so that  $X$  stays on  $M$  for all time. Thus from here on we are considering the case of Brownian motion with non-explosive drift on a complete Riemannian manifold.

We note here that all the results in Sect. 1.1 carry over to the manifold setting with  $(M, \text{dist})$  being the Riemannian manifold (with the distance  $\text{dist}$  induced by the Riemannian metric) and  $m$  taken to be the volume measure.

### 3.3 The interface

Varadhan small-time asymptotics and Lemma 3 can be used to show the following: that the existence of an MMC implies that, for each time  $t$ , there is a *deterministic involutive isometry*  $F_t$  which exchanges  $X_t$  with  $Y_t$  and fixes the set of points equidistant from both  $X_t$  and  $Y_t$ . This generalizes the time-varying reflection isometry of Euclidean space which is mentioned in Remark 11; the fixed-point set of  $F_t$  corresponds to the ‘evolving mirror’ of the Euclidean case.

The rôle of Varadhan’s small-time asymptotics in the following is analogous to the rôle of Lemma 7 in the Euclidean case. This powerful technique gives the logarithmic asymptotics of the density of  $X_t$  when  $t \downarrow 0$ , as stated in the following lemma.

**Lemma 31** *Suppose that  $X$  satisfies the assumptions of both diffusion-geodesic completeness and stochastic completeness. Let  $M_1$  and  $M_2$  be compact subsets of  $M$ . Then the density  $p$  of  $X_t$  satisfies the following:*

$$\lim_{t \downarrow 0} 2t \log p(\mathbf{x}; t, \mathbf{y}) = -\text{dist}^2(\mathbf{x}, \mathbf{y}) \quad (58)$$

*uniformly for all  $\mathbf{x}, \mathbf{y} \in M_1 \times M_2$ , where  $\text{dist}(\mathbf{x}, \mathbf{y})$  is the Riemannian distance between  $\mathbf{x}$  and  $\mathbf{y}$ .*

This theorem was proven by [39] for diffusion processes on Euclidean space. Later [29] noticed that Varadhan’s arguments carry over to diffusions on closed manifolds whose generators are of the form  $L = \frac{1}{2} \Delta_M + \mathbf{b}$ . Molchanov also showed that this result could be extended to general smooth complete manifolds by introducing a reflected diffusion in a suitably large domain  $U \subset M$  containing  $\mathbf{x}$  and  $\mathbf{y}$ , with the same generator  $L$  inside, and using this process to define a natural diffusion on the ‘double’  $U$ . He then showed that smoothing techniques allowed the approximation of the ‘double’  $U$  by a smooth closed manifold, such that the diffusion thus defined has a density that is sufficiently close to that of the original one [29, p. 18 and further references].

We can now restate the pivotal Theorem 10 from Sect. 2.1 in the new context of manifolds. The proof of the manifold case follows that of the Euclidean case, but uses Lemma 31 in place of Lemma 7, and uses the strong maximum principle (Lemma 9) in local coordinates; we omit details.



**Theorem 32** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. For any  $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}(\mu_s)$ , and any  $s > 0$ , the following equalities hold:*

$$\begin{aligned} I(\mathbf{x}_0, \mathbf{y}_0, s) &= H(\mathbf{x}, \mathbf{y}), \\ I^-(\mathbf{x}_0, \mathbf{y}_0, s) &= H^-(\mathbf{x}, \mathbf{y}), \\ I^+(\mathbf{x}_0, \mathbf{y}_0, s) &= H^+(\mathbf{x}, \mathbf{y}). \end{aligned}$$

Let  $\tau' = \inf\{s > 0 : X_s \in I(\mathbf{x}_0, \mathbf{y}_0, s)\}$  be the first time that  $X$  hits the interface. Then the following holds.

**Corollary 33** *Almost surely  $\tau' = \tau$ , so coupling occurs when  $X$  first hits the interface. Furthermore,  $\mu$ -almost surely, for all  $t < \tau$ ,*

$$\begin{aligned} I(\mathbf{x}_0, \mathbf{y}_0, t) &= H(X_t, Y_t), \quad I^-(\mathbf{x}_0, \mathbf{y}_0, t) = H^-(X_t, Y_t), \\ I^+(\mathbf{x}_0, \mathbf{y}_0, t) &= H^+(X_t, Y_t). \end{aligned} \quad (59)$$

*Proof* The proof follows the lines of the proof of Corollary 12. The only additional detail that we have to check here (which was immediate in the Euclidean case) is that, for any  $t > 0$  with  $X_t \neq Y_t$ , any  $\mathbf{z} \in H(X_t, Y_t)$  and any rational sequence  $t_n \downarrow t$ , there is  $\mathbf{z}_n \in H(X_{t_n}, Y_{t_n})$  such that  $\mathbf{z}_n \rightarrow \mathbf{z}$ . This was used in Corollary 12 to show  $H(X_t, Y_t) \subseteq I(\mathbf{x}_0, \mathbf{y}_0, t)$ .

Recall the event  $E = \cap_{q \in \mathcal{Q}} E_q$ , where  $E_q$  was defined in (19). Assume  $E$  holds. For notational convenience, denote  $H(X_t, Y_t)$ ,  $X_t, Y_t$  by  $H, \mathbf{x}, \mathbf{y}$  and  $H(X_{t_n}, Y_{t_n})$ ,  $X_{t_n}, Y_{t_n}$  by  $H_n, \mathbf{x}_n, \mathbf{y}_n$  respectively. Let  $\gamma : [0, 2 \operatorname{dist}(\mathbf{x}, \mathbf{z})] \rightarrow M$  denote the continuous curve such that  $\gamma|_{[0, \operatorname{dist}(\mathbf{x}, \mathbf{z})]}$  is a minimal geodesic joining  $\mathbf{x}$  and  $\mathbf{z}$  and  $\gamma|_{[\operatorname{dist}(\mathbf{x}, \mathbf{z}), 2 \operatorname{dist}(\mathbf{x}, \mathbf{z})]}$  is a minimal geodesic joining  $\mathbf{z}$  and  $\mathbf{y}$ . As  $M$  has no branching geodesics, it follows that  $\operatorname{dist}(\mathbf{x}, \gamma(s)) < \operatorname{dist}(\mathbf{y}, \gamma(s))$  for any  $s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z})]$ . Consequently for any  $\delta > 0$ , by the compactness of  $\{\gamma(s) : s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z}) - \delta]\}$ ,  $\min_{s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z}) - \delta]} (\operatorname{dist}(\mathbf{y}, \gamma(s)) - \operatorname{dist}(\mathbf{x}, \gamma(s))) > 0$  and hence,  $\min_{s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z}) - \delta]} (\operatorname{dist}(\mathbf{y}_n, \gamma(s)) - \operatorname{dist}(\mathbf{x}_n, \gamma(s))) > 0$  for sufficiently large  $n$ . Thus, for sufficiently large  $n$ ,  $\gamma(s) \in H^-(\mathbf{x}_n, \mathbf{y}_n) = I^-(\mathbf{x}_0, \mathbf{y}_0, t_n)$  for all  $s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z}) - \delta]$  and consequently,  $\min_{s \in [0, \operatorname{dist}(\mathbf{x}, \mathbf{z}) - \delta]} \alpha(t_n, \gamma(s)) > 0$ . Similarly,  $\min_{s \in [\operatorname{dist}(\mathbf{x}, \mathbf{z}) + \delta, 2 \operatorname{dist}(\mathbf{x}, \mathbf{z})]} \alpha(t_n, \gamma(s)) < 0$  for sufficiently large  $n$ . Thus, as  $E$  holds, the continuity of  $\alpha(t_n, \cdot)$ , implies that for sufficiently large  $n$ , there is  $\mathbf{z}_n \in \gamma \cap H_n$  such that  $\mathbf{z}_n \rightarrow \mathbf{z}$ . As  $\mu(E) = 1$ , this implies  $H(X_t, Y_t) \subseteq I(\mathbf{x}_0, \mathbf{y}_0, t)$  almost surely.

The rest of the proof carries over verbatim from that of Corollary 12.  $\square$

The striking fact that emerges from the above is that, almost surely under the coupling  $\mu$ , for each  $s > 0$ ,  $H(X_t, Y_t)$  is a *non-random set which depends only on  $s$  and not on the specific location of  $(X_t, Y_t)$* . We will call this set  $H_t$  henceforth. Similarly, denote  $H_t^+ = H^+(X_t, Y_t)$  and  $H_t^- = H^-(X_t, Y_t)$ . The family  $\{H_t : t \geq 0\}$  corresponds to the family of *moving mirrors* from Sect. 2.

We now follow [25]'s construction to define a deterministic global involutive isometry  $F_s$  which fixes  $H_s$  and maps  $X_s$  to  $Y_s$  under the coupling. The argument of [25, Lemma 4.6] applies directly to our case: we therefore omit proof.

**Lemma 34** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Take  $s \geq 0$ . If  $\mathbf{x}, \mathbf{y} \in M$ , with  $\mathbf{x} \neq \mathbf{y}$ , satisfies*

$$\text{dist}(\mathbf{x}, \mathbf{z}) = \text{dist}(\mathbf{y}, \mathbf{z}) \quad (60)$$

*for all  $\mathbf{z} \in H_s$ , then  $(\mathbf{x}, \mathbf{y}) \in H_s^+ \times H_s^- \cup H_s^- \times H_s^+$  (so  $\mathbf{x}$  and  $\mathbf{y}$  lie in opposite “half-manifolds”). Furthermore, for any  $\mathbf{x} \in M$ , a point  $\mathbf{y} \in M \setminus \{\mathbf{x}\}$  satisfying (60) is unique if it exists.*

Whenever such a  $\mathbf{y}$  exists, we will call  $\mathbf{y}$  the *mirror image* of  $\mathbf{x}$  at time  $s$ . With the aid of the above lemma, the isometry  $F_s$  is constructed using a procedure which is similar to [25, Theorem 4.5], but is subject to some modification as described in the following lemma and its proof.

**Lemma 35** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Assume  $(X, Y)$  is a Markovian maximal coupling with starting points  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . Then, for each  $s \in [0, \tau]$ , there is a deterministic involutive isometry  $F_s$  with fixed point set  $H_s$  such that  $Y_s = F_s(X_s)$ , furthermore  $F_s(H_s^-) = H_s^+$ .*

*Proof* Define the set

$$A_s = \{\mathbf{x} \in M : \text{there exists } \mathbf{y} \in M \setminus \{\mathbf{x}\} \text{ such that (60) holds}\}.$$

For  $\mathbf{x} \in A_s$ , define  $F_s(\mathbf{x})$  to be the unique  $\mathbf{y}$  for which (60) holds. For  $\mathbf{x} \in H_s$ , define  $F_s(\mathbf{x}) = \mathbf{x}$ . Following the proof of [25, Theorem 4.5], the set  $\hat{A}_s = A_s \cup H_s$  is closed. Furthermore, by Theorem 32 and Lemma 2, on the event  $[0 < s < \tau]$  the support of  $X_s$  (equivalently  $Y_s$ ) is the whole of  $H_s^-$  (respectively  $H_s^+$ ). This, by Lemma 3 and Theorem 32, implies  $\hat{A}_s = M$  for all  $s > 0$ .

A little more argument is required for  $s = 0$ . By Theorem 32, Lemmas 31 and 2, for any  $\mathbf{x} \in H_0^-$ , there is a sequence  $t_n \downarrow 0$  and  $\mathbf{x}_n \rightarrow \mathbf{x}$  such that  $\mathbf{x}_n \in A_{t_n}$  with  $\mathbf{y}_n \in M$  being its *mirror image* at time  $t_n$ , for all  $n$ . Take any  $\mathbf{z}_0 \in H_0$ . Following the proof of Corollary 33, for sufficiently large  $n$ , there is  $\mathbf{z}_n \in H_{t_n}$  such that  $\mathbf{z}_n \rightarrow \mathbf{z}_0$ . As  $\text{dist}(\mathbf{x}_n, \mathbf{z}_n) = \text{dist}(\mathbf{y}_n, \mathbf{z}_n)$ , it follows that the set of distances  $\{\text{dist}(\mathbf{z}_0, \mathbf{y}_n)\}_{n \geq 1}$  is bounded. Consequently the properness of  $M$  implies that there is a subsequence  $\{n_k\}$  such that  $\mathbf{y}_{n_k} \rightarrow \mathbf{y}$  for some  $\mathbf{y} \in M$ . Now, for any  $\mathbf{z} \in H_0$ , take  $\mathbf{z}'_n \in H_{t_n}$  such that  $\mathbf{z}'_n \rightarrow \mathbf{z}$ . Thus,

$$\begin{aligned} \text{dist}(\mathbf{y}, \mathbf{z}) &= \lim_{k \rightarrow \infty} \text{dist}(\mathbf{y}_{n_k}, \mathbf{z}'_{n_k}) \\ &= \lim_{k \rightarrow \infty} \text{dist}(\mathbf{x}_{n_k}, \mathbf{z}'_{n_k}) = \text{dist}(\mathbf{x}, \mathbf{z}). \end{aligned}$$

This implies  $\hat{A}_0 = M$ . Note that, by Lemma 34, the limit  $\mathbf{y}$  is uniquely determined by  $\mathbf{x}$  and  $H_0$ , and thus, does not depend on the subsequence chosen. This implies  $\mathbf{y}_n \rightarrow \mathbf{y}$ . Define  $F_0(\mathbf{x}) = \mathbf{y}$ .

Thus  $F_s$  is defined on the whole of  $M$  for every  $s \geq 0$ . Continuity of  $F_s$  for  $s \geq 0$  follows exactly along the lines of the proof of continuity of the map  $R$  in [25, Theorem 4.5]. Further, by definition,  $F_s$  is involutive. Thus, in particular,  $F_s$  is an open map.

To prove that  $F_s$  is, in fact, an isometry, we have to modify the proof of [25, Lemma 5.3] appropriately, as we outline in the following.

First, consider  $s > 0$ . If  $\mathbf{x}, \mathbf{y} \in H_s$  or  $\mathbf{x} \in H_s^-, \mathbf{y} \in H_s^+$ , then  $\text{dist}(\mathbf{x}, \mathbf{y}) = \text{dist}(F_s(\mathbf{x}), F_s(\mathbf{y}))$  follows from the definition of  $F_s$ . So, assume  $\mathbf{x}, \mathbf{y} \in H_s^-$ . Take  $\delta > 0$  small enough such that

$$\overline{\mathcal{B}(\mathbf{x}, \delta)} \subset H_s^-, \overline{\mathcal{B}(\mathbf{y}, \delta)} \subset H_s^-, \overline{\mathcal{B}(F_s(\mathbf{x}), \delta)} \subset H_s^+, \overline{\mathcal{B}(F_s(\mathbf{y}), \delta)} \subset H_s^+.$$

Let

$$V_1 = \mathcal{B}(\mathbf{x}, \delta) \cap F_s(\mathcal{B}(F_s(\mathbf{x}), \delta)), \quad V_2 = \mathcal{B}(\mathbf{y}, \delta) \cap F_s(\mathcal{B}(F_s(\mathbf{y}), \delta)), \quad U_2 = \mathcal{B}(\mathbf{y}, \delta/2) \cap F_s(\mathcal{B}(F_s(\mathbf{y}), \delta/2)).$$

For  $t > 0$ , by the strong Markov property, Corollary 33 and Lemma 2, we have

$$\begin{aligned} & \mu(X_{s+t} \in U_2, X_s \in V_1, \tau > s+t) \\ &= \int_{V_1} \alpha^+(s, \mathbf{z}) \left\{ \int_{U_2} (p(\mathbf{z}; t, \mathbf{w}) - p(F_s(\mathbf{z}); t, \mathbf{w})) m(d\mathbf{w}) \right\} m(d\mathbf{z}). \end{aligned} \quad (61)$$

Similarly,

$$\begin{aligned} & \mu(Y_{s+t} \in F_{s+t}(U_2), Y_s \in F_s(V_1), \tau > s+t) \\ &= \int_{F_s(V_1)} \alpha^-(s, \mathbf{z}) \left\{ \int_{F_{s+t}(U_2)} (p(\mathbf{z}; t, \mathbf{w}) - p(F_s(\mathbf{z}); t, \mathbf{w})) m(d\mathbf{w}) \right\} m(d\mathbf{z}). \end{aligned} \quad (62)$$

Observe that if  $\mathbf{z}, \mathbf{w} \in H_s^-$  or  $\mathbf{z}, \mathbf{w} \in H_s^+$ , then  $\text{dist}(\mathbf{z}, \mathbf{w}) < \text{dist}(F_s(\mathbf{z}), \mathbf{w})$ . To see this, let  $\gamma$  be the minimal geodesic joining  $\mathbf{w}$  and  $F_s(\mathbf{z})$  and let  $\mathbf{z}_0 \in \gamma \cap H_s$ . Then

$$\text{dist}(\mathbf{z}, \mathbf{w}) \leq \text{dist}(\mathbf{z}, \mathbf{z}_0) + \text{dist}(\mathbf{z}_0, \mathbf{w}) = \text{dist}(F_s(\mathbf{z}), \mathbf{z}_0) + \text{dist}(\mathbf{z}_0, \mathbf{w}) = \text{dist}(F_s(\mathbf{z}), \mathbf{w}).$$

If equality holds in the first inequality above, then we can take a minimal geodesic joining  $\mathbf{z}$  and  $\mathbf{w}$  that branches from  $\gamma$  at  $\mathbf{z}_0$  which gives a contradiction.

Next, we claim that there is  $\epsilon > 0$  such that for  $t \in [0, \epsilon]$ ,  $F_{t+s}(U_2) \subseteq F_s(V_2)$ . Suppose not. Then there is a sequence  $t_n \downarrow 0$  and  $\mathbf{x}_n \in U_2$  such that  $\mathbf{y}_n = F_{s+t_n}(\mathbf{x}_n) \in F_s(V_2^c)$ . As  $U_2$  is bounded, we obtain a subsequence  $n_k$  such that  $\mathbf{x}_{n_k} \rightarrow \mathbf{x}^o \in \bar{U}_2$  as  $k \rightarrow \infty$ . Take any  $\mathbf{z}^o \in H_s$ . Following the proof of Corollary 33, for sufficiently large  $n$ , there is  $\mathbf{z}_n^o \in H_{s+t_n}$  such that  $\mathbf{z}_n^o \rightarrow \mathbf{z}^o$ . As  $\text{dist}(\mathbf{x}_{n_k}, \mathbf{z}_{n_k}^o) = \text{dist}(\mathbf{y}_{n_k}, \mathbf{z}_{n_k}^o)$ ,

$$\begin{aligned} \text{dist}(\mathbf{y}_{n_k}, \mathbf{z}^o) &\leq \text{dist}(\mathbf{y}_{n_k}, \mathbf{z}_{n_k}^o) + \text{dist}(\mathbf{z}_{n_k}^o, \mathbf{z}^o) = \text{dist}(\mathbf{x}_{n_k}, \mathbf{z}_{n_k}^o) + \text{dist}(\mathbf{z}_{n_k}^o, \mathbf{z}^o) \\ &\leq \text{dist}(\mathbf{x}_{n_k}, \mathbf{x}^o) + \text{dist}(\mathbf{x}^o, \mathbf{z}^o) + 2 \text{dist}(\mathbf{z}_{n_k}^o, \mathbf{z}^o). \end{aligned}$$

Thus,  $\mathbf{y}_{n_k}$  is bounded and we can extract a further subsequence  $n_{k_l}$  such that  $\mathbf{y}_{n_{k_l}} \rightarrow \mathbf{y}^o$  as  $l \rightarrow \infty$ . As  $F_s$  is a bijective open map,  $F_s(V_2^c)$  is closed and hence,  $\mathbf{y}^o \in F_s(V_2^c)$ . Now, take any  $\mathbf{z} \in H_s$ . Taking a sequence  $\mathbf{z}_{n_{k_l}} \in H_{s+t_{n_{k_l}}}$  such that  $\mathbf{z}_{n_{k_l}} \rightarrow \mathbf{z}$ , we observe

$$\text{dist}(\mathbf{x}^o, \mathbf{z}) = \lim_{l \rightarrow \infty} \text{dist}(\mathbf{x}_{n_{k_l}}, \mathbf{z}_{n_{k_l}}) = \lim_{l \rightarrow \infty} \text{dist}(\mathbf{y}_{n_{k_l}}, \mathbf{z}_{n_{k_l}}) = \text{dist}(\mathbf{y}^o, \mathbf{z}).$$

By Lemma 34,  $\mathbf{y}^o = F_s(\mathbf{x}^o)$ , which gives a contradiction as  $\mathbf{x}^o \in \overline{U_2} \subseteq V_2$  but  $\mathbf{y}^o \in F_s(V_2^c)$ . The claim follows from this.

The above two observations along with Lemma 31 applied to (61) and (62) yield

$$\begin{aligned} & \lim_{t \downarrow 0} 2t \log [\mu(X_{s+t} \in U_2, X_s \in V_1, \tau > s+t)] \\ &= - \inf_{\mathbf{z} \in V_1, \mathbf{w} \in U_2} \text{dist}^2(\mathbf{z}, \mathbf{w}), \\ & \limsup_{t \downarrow 0} 2t \log [\mu(Y_{s+t} \in F_{s+t}(U_2), X_s \in F_s(V_1), \tau > s+t)] \\ &\leq - \inf_{\mathbf{z} \in V_1, \mathbf{w} \in V_2} \text{dist}^2(F_s(\mathbf{z}), F_s(\mathbf{w})). \end{aligned}$$

Since the left hand side of (61) is the same as that of (62), we take  $\delta \downarrow 0$  above to get

$$\text{dist}(\mathbf{x}, \mathbf{y}) \geq \text{dist}(F_s(\mathbf{x}), F_s(\mathbf{y})).$$

As  $F_s$  is involutive, applying a symmetric argument with  $\mathbf{x}, \mathbf{y}$  replaced by  $F_s(\mathbf{x}), F_s(\mathbf{y})$  yield the opposite inequality. Hence,  $\text{dist}(\mathbf{x}, \mathbf{y}) = \text{dist}(F_s(\mathbf{x}), F_s(\mathbf{y}))$  for all  $\mathbf{x}, \mathbf{y} \in M$ . Thus,  $F_s$  is an isometry for every  $s > 0$ .

Finally, consider the case  $s = 0$ . Again, for  $\mathbf{x}, \mathbf{y} \in H_0$  or  $\mathbf{x} \in H_0^-, \mathbf{y} \in H_0^+$ ,  $\text{dist}(\mathbf{x}, \mathbf{y}) = \text{dist}(F_0(\mathbf{x}), F_0(\mathbf{y}))$  follows from the definition of  $F_0$ . For  $\mathbf{x}, \mathbf{y} \in H_0^-$ , by the same procedure used to define  $F_0$  earlier in the proof, we obtain sequences  $t_n \downarrow 0$  and  $\mathbf{x}_n \in H_{t_n}^-$  and  $\mathbf{y}_n \in H_{t_n}^+$  such that  $\mathbf{x}_n \rightarrow \mathbf{x}, \mathbf{y}_n \rightarrow \mathbf{y}, F_{t_n}(\mathbf{x}_n) \rightarrow F_0(\mathbf{x})$  and  $F_{t_n}(\mathbf{y}_n) \rightarrow F_0(\mathbf{y})$ . Thus,

$$\text{dist}(F_0(\mathbf{x}), F_0(\mathbf{y})) = \lim_{n \rightarrow \infty} \text{dist}(F_{t_n}(\mathbf{x}_n), F_{t_n}(\mathbf{y}_n)) = \lim_{n \rightarrow \infty} \text{dist}(\mathbf{x}_n, \mathbf{y}_n) = \text{dist}(\mathbf{x}, \mathbf{y}),$$

which proves that  $F_0$  is an isometry.

Now,  $F_s(H_s^-) = H_s^+$  follows from Lemma 34. This completes the proof of the lemma.  $\square$

Following [32, Chapter 10, Proposition 24], as  $H_s$  is the fixed point set of an isometry therefore each connected component of  $H_s$  is a totally geodesic submanifold (in particular, a smooth submanifold). Furthermore, as  $H_s$  partitions  $M$  into two disjoint open subsets, it can be verified (for example by referring to normal coordinates based around a point in  $H_s$ ) that  $H_s$  must be of codimension 1. Furthermore, this discussion also implies that for any  $\mathbf{x}, \mathbf{y} \in M$  there is *at most one isometry* whose set of fixed points is the set  $H(\mathbf{x}, \mathbf{y})$ . We will refer to this isometry, if it exists, as  $f_{\mathbf{x}, \mathbf{y}}$ . In fact

Lemmas 34 and 35 together imply that for any  $s \geq 0$  there does indeed exist such a  $f_{\mathbf{x},\mathbf{y}}$  for each  $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}(\mu_s)$ , given by

$$f_{\mathbf{x},\mathbf{y}} = F_s.$$

To get an intuitive picture of how  $F_s$  acts locally around a point  $\mathbf{x}^* \in H_s$  (hence, fixed by  $F_s$ ), recall that

$$d F_s : T_{\mathbf{x}^*} M \rightarrow T_{\mathbf{x}^*} M$$

is a linear isometry. We can form an orthonormal basis  $e_1, \dots, e_d$  of  $T_{\mathbf{x}^*} M$  such that  $e_1, \dots, e_{d-1}$  form a basis of the tangent space  $T_{\mathbf{x}^*} H_s$  viewed as a subspace of  $T_{\mathbf{x}^*} M$ . Because  $H_s$  is totally geodesic, these vectors correspond to geodesics through  $\mathbf{x}^*$  that stay in  $H_s$ . As  $H_s$  is the fixed point set of  $F_s$ , the basis vectors  $e_1, \dots, e_{d-1}$  must be fixed by  $d F_s$ , while  $e_d$  is mapped by  $d F_s$  to  $-e_d$ . Thus, locally, one geodesic passing through  $\mathbf{x}^*$  is inverted by  $F_s$ , while geodesics starting in directions orthogonal to the inverted geodesic are fixed by  $F_s$ .

### 3.4 Structure of the manifold $M$

In this section, we will use the isometries  $f_{\mathbf{x},\mathbf{y}}$  constructed above for every pair of points  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  to show that the underlying complete Riemannian manifold  $M$  is *homogeneous* (i.e. the isometry group acts transitively) and *isotropic* about a chosen point  $\mathbf{x}^*$  (i.e. there are  $\frac{d(d-1)}{2}$  independent rotations about  $\mathbf{x}^*$ ). This will imply that  $M$  is a *maximally symmetric space*, i.e. the isometry group  $\mathcal{G}$  of  $M$  has the maximal dimension possible (namely,  $\frac{d(d+1)}{2}$ ) for any  $d$ -dimensional manifold. It is an almost immediate consequence that the space  $M$  can be classified (up to scaling) as one of the three model space forms of constant curvatures respectively  $-1, 0$ , and  $+1$ .

**Lemma 36** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Under LPC,  $(M, g)$  is a homogeneous space.*

*Proof* We want to show that  $\mathcal{G}$  acts transitively on  $M$ . Together with LPC, the work of the previous subsection shows that for each  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$ , there exists an involutive isometry  $f_{\mathbf{x},\mathbf{y}}$ . This implies that, for any  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$ , there is an isometry  $G_{\mathbf{x}_0,\mathbf{x}} = f_{\mathbf{y}_0,\mathbf{x}} \circ f_{\mathbf{x}_0,\mathbf{y}_0}$  which takes  $\mathbf{x}_0$  to  $\mathbf{x}$ . Consider the set of isometries

$$\mathcal{I} = \{G_{\mathbf{x}_0,\mathbf{x}} : \mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)\}.$$

Let  $\mathcal{H}$  be the closure of the subgroup generated by  $\mathcal{I}$ , so  $\mathcal{H}$  is a closed subgroup of  $\mathcal{G}$ . Denote by  $\mathcal{O}(\mathbf{x}_0)$ , the *orbit* or *set of equivalent points* of  $\mathbf{x}_0$  under  $\mathcal{H}$ . By construction,  $\mathcal{B}(\mathbf{x}_0, r) \subseteq \mathcal{O}(\mathbf{x}_0)$ . In order to prove that  $M$  is homogeneous, we need to prove  $\mathcal{O}(\mathbf{x}_0) = M$ , which we will show by proving that  $\mathcal{O}(\mathbf{x}_0)$  is both open and closed in  $M$ . Let  $\mathbf{z}$  be a limit point of  $\mathcal{O}(\mathbf{x}_0)$ . Then, there is a sequence of isometries  $G_n \in \mathcal{H}$  such that  $G_n(\mathbf{x}_0) \rightarrow \mathbf{z}$ . By [30, p. 7], there exists an isometry  $G \in \mathcal{H}$  and a subsequence  $G_{n_k} \in \mathcal{H}$  such that  $G_{n_k} \rightarrow G$  in the topology of isometries (i.e.  $G_{n_k}(\mathbf{x}) \rightarrow G(\mathbf{x})$

for all  $\mathbf{x} \in M$ ), and consequently,  $G(\mathbf{x}_0) = \mathbf{z}$ . This shows that  $\mathcal{O}(\mathbf{x}_0)$  is closed. On the other hand, if  $\mathbf{y} \in \mathcal{O}(\mathbf{x}_0)$ , then there is an isometry  $G \in \mathcal{H}$  such that  $\mathbf{y} = G(\mathbf{x}_0)$ . Therefore,  $\mathcal{B}(\mathbf{y}, r) = G(\mathcal{B}(\mathbf{x}_0, r)) \subseteq \mathcal{O}(\mathbf{x}_0)$  (as  $\mathcal{B}(\mathbf{x}_0, r) \subseteq \mathcal{O}(\mathbf{x}_0)$ ) implying  $\mathcal{O}(\mathbf{x}_0)$  is open. Thus,  $\mathcal{O}(\mathbf{x}_0) = M$ , proving the lemma.  $\square$

In the following lemma, we will write  $\mathbf{x}^*$  for the midpoint of a minimal geodesic  $\gamma_{\mathbf{x}_0, \mathbf{y}_0}$  connecting  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . If two vectors  $u, v$  belong to the same tangent space then we denote the angle between them by  $\angle(u, v)$ .

**Lemma 37** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Under LPC,  $M$  is isotropic at  $\mathbf{x}^*$ .*

*Proof* Let  $\gamma(v)$  denote the geodesic issuing from  $\mathbf{x}^*$  in direction  $v$ . Suppose  $\gamma(v_0) = \gamma_{\mathbf{x}_0, \mathbf{y}_0}$ , thus defining a unit vector  $v_0$ . The proof proceeds in three steps as follows.

**Step 1.** First, we want to show that there is  $\varepsilon > 0$  such that, for any  $v \in T_{\mathbf{x}^*}M$  with  $\angle(v, v_0) < \varepsilon$ , there is an isometry  $g_v$  leaving  $\mathbf{x}^*$  fixed and  $dg_v(v_0) = v$ .

By continuity of geodesics in the starting direction, we can choose  $\varepsilon > 0$  sufficiently small so that  $\gamma(v')$  intersects  $\mathcal{B}(\mathbf{x}_0, r)$  and  $\gamma(-v')$  intersects  $\mathcal{B}(\mathbf{y}_0, r)$  whenever  $\angle(v', v_0) < \varepsilon$ . By [32, Proposition 20, p. 141], with a possibly smaller choice of  $\varepsilon > 0$ , we can take  $\mathbf{x}_{v'} \in \gamma(v') \cap \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y}_{v'} \in \gamma(-v') \cap \mathcal{B}(\mathbf{y}_0, r)$  such that  $\gamma(v')$  realises the distance  $\text{dist}(\mathbf{x}^*, \mathbf{x}_{v'})$  and  $\gamma(-v')$  realises the distance  $\text{dist}(\mathbf{x}^*, \mathbf{y}_{v'})$ . Furthermore, by continuity of the metric, when  $\varepsilon > 0$  is small enough, we can take such  $\mathbf{x}_{v'}, \mathbf{y}_{v'}$  satisfying  $\text{dist}(\mathbf{x}_{v'}, \mathbf{x}^*) = \text{dist}(\mathbf{y}_{v'}, \mathbf{x}^*)$  whenever  $\angle(v', v_0) < \varepsilon$ . Thus, from the developments of the previous subsection, there is an involutive isometry  $f_{\mathbf{x}_{v'}, \mathbf{y}_{v'}}$  which fixes  $\mathbf{x}^*$ , inverts the geodesic passing through  $\mathbf{x}^*$  in direction  $v'$ , and fixes all the geodesics which pass through  $\mathbf{x}^*$  in directions orthogonal to  $v'$ .

Now, take any unit vector  $v \in T_{\mathbf{x}^*}M$  with  $\angle(v, v_0) < 2\varepsilon$ . Let  $v' = \frac{v+v_0}{|v+v_0|}$ . By the properties of rhombuses,  $\angle(v', v_0) = \frac{1}{2}\angle(v, v_0) < \varepsilon$ , and thus  $f_{\mathbf{x}_{v'}, \mathbf{y}_{v'}}$  exists as specified in the preceding paragraph. Now, consider the isometry  $g_v = f_{\mathbf{x}_{v'}, \mathbf{y}_{v'}} \circ f_{\mathbf{x}_0, \mathbf{y}_0}$ . Note that  $g_v$  fixes  $\mathbf{x}^*$  and a straightforward calculation reveals  $dg_v(v_0) = v$ . This  $g_v$  is our required isometry.

**Step 2.** Take any unit vector  $w \in T_{\mathbf{x}^*}M$  such that  $w$  and  $v_0$  are linearly independent. Let  $\Pi$  be the two-dimensional subspace of  $T_{\mathbf{x}^*}M$  generated by  $v_0$  and  $w$  and denote by  $\mathbb{S}(v_0, w)$  the circle in  $T_{\mathbf{x}^*}M$  centred at the origin of  $T_{\mathbf{x}^*}M$  and running through  $v_0$  and  $w$ . Let  $U$  be a normal neighbourhood around  $\mathbf{x}^*$ . Let  $S_\Pi = \exp_{\mathbf{x}^*}(\Pi) \cap U$  denote the two-dimensional fragment of  $M$  corresponding to  $\Pi$  and lying in  $U$ .

Denote by  $\mathcal{H}(v_0, w)$  the closed subgroup of isometries generated by  $\{g_v : v \in \mathbb{S}(v_0, w)\}$ ,

$\angle(v, v_0) < \varepsilon$ , where  $g_v$  are the isometries constructed in Step 1. Note that the set  $\{g_v : v \in \mathbb{S}(v_0, w)\}$ ,

$\angle(v, v_0) < \varepsilon$ , and hence  $\mathcal{H}(v_0, w)$ , fixes  $\mathbf{x}^*$  and keeps vectors orthogonal to  $\{v_0, w\}$  fixed. Let

$$O(v_0) = \{dg(v_0) : g \in \mathcal{H}(v_0, w)\}.$$

We want to show that  $O(v_0) = \mathbb{S}(v_0, w)$ .

Note that, if  $v_n = d g_n(v_0)$  such that  $v_n \rightarrow v$ , then, by the fact that  $g_n(\mathbf{x}^*) = \mathbf{x}^*$  for all  $n$ , we can choose a subsequence  $g_{n_k}$  and a  $g \in \mathcal{H}(v_0, w)$  such that  $g_{n_k} \rightarrow g$  in the topology of isometries [30, p. 7]. Thus, by [30, Lemma 4],  $d g_{n_k}(v_0) \rightarrow d g(v_0)$  implying  $O(v_0)$  is closed. Furthermore, if  $g \in \mathcal{H}(v_0, w)$  then  $d g$  is a linear isometry on  $T_{\mathbf{x}^*}M$ . So the same argument as in the previous lemma shows that  $O(v_0)$  is open. Thus,  $O(v_0) = \mathbb{S}(v_0, w)$ .

Thus, in particular, the subgroup of isometries  $\mathcal{G}_{\mathbf{x}^*}$  which fix  $\mathbf{x}^*$  (the *isotropy group* at  $\mathbf{x}^*$ ) generates all the rotations of  $T_{\mathbf{x}^*}M$  based at  $\mathbf{x}^*$  in 2-planes containing  $v_0$ . We describe the isometries in  $\mathcal{H}(v_0, w)$  as *rotations* in  $\mathbb{S}(v_0, w)$ .

**Step 3.** We will now show that, given two ordered orthonormal frames based at  $T_{\mathbf{x}^*}M$ , there is a sequence of isometries in  $\mathcal{G}_{\mathbf{x}^*}$  that take one to the other. In particular this implies that  $M$  is isotropic at  $\mathbf{x}^*$ . Let  $(e_1, \dots, e_d)$  and  $(e'_1, \dots, e'_d)$  be ordered orthonormal frames in  $T_{\mathbf{x}^*}M$ . We can apply rotations in  $\mathbb{S}(v_0, e_1)$  (respectively  $\mathbb{S}(v_0, e'_d)$ ) to align  $e_1$  with  $v_0$  (respectively  $e'_d$  with  $v_0$ ). Thus, without loss of generality, we consider frames of the form  $(v_0, e_2, \dots, e_d)$  and  $(e'_1, \dots, e'_{d-1}, v_0)$ .

Now, apply a rotation in  $\mathbb{S}(v_0, e'_1)$  to transform  $(v_0, e_2, \dots, e_d)$  to  $(e'_1, e_2^{(1)}, \dots, e_d^{(1)})$  for some unit vectors  $e_2^{(1)}, \dots, e_d^{(1)}$  in  $T_{\mathbf{x}^*}M$ . If  $v_0$  and  $e_2^{(1)}$  are linearly independent, then apply a rotation in  $\mathbb{S}(v_0, e_2^{(1)})$ , to bring  $(e'_1, e_2^{(1)}, \dots, e_d^{(1)})$  to  $(e'_1, v_0, e_3^{(2)}, \dots, e_d^{(2)})$ . If  $e_2^{(1)} = -v_0$ , then achieve the same result using the reflection  $f_{\mathbf{x}_0, v_0}$ . Note that these operations both keep  $e'_1$  fixed as it is orthogonal to  $\{v_0, e_2^{(1)}\}$ .

The same procedure is applied inductively to  $(e'_1, v_0, e_3^{(2)}, \dots, e_d^{(2)})$  to obtain  $(e'_1, e'_2, v_0, e_4^{(3)}, \dots, e_d^{(4)})$  (note that these operations leave  $e'_1$  fixed), and so on. Finally we obtain  $(e'_1, \dots, e'_{d-1}, v_0)$ , which proves the lemma.  $\square$

The above two lemmas imply the following rigidity theorem which completely classifies the space  $M$ .

**Theorem 38** *Suppose that the complete, connected Riemannian manifold  $M$  supports Brownian motion with drift for which there is a Markovian maximal coupling and moreover LPC holds. Then  $M$  has constant sectional curvature. Moreover  $M$  must be simply connected and therefore (up to scaling)  $M$  must be one of the three model spaces  $\mathbb{R}^d$ ,  $\mathbb{S}^d$  and  $\mathbb{H}^d$ .*

*Proof* By Lemmas 36 and 37, we see that  $M$  is a *maximally symmetric space*, i.e., the dimension of  $\text{Iso}(M)$  is  $\frac{d(d+1)}{2}$  [36, p. 195]. In particular, this implies that  $M$  has constant sectional curvature [32, p. 190]. For the second part of the corollary, the argument of [32, p. 190] shows that a complete, connected maximally symmetric Riemannian manifold must be one of the three model spaces above, or  $\mathbb{R}\mathbb{P}^d$ . But, as observed in [25, Example 6.4], there is no involutive isometry of  $\mathbb{R}\mathbb{P}^d$  of the form described in Lemma 35. This proves the theorem.  $\square$

**Remark 39** For the three model spaces described above, for every  $\mathbf{x}, \mathbf{y} \in M$ , the reflection isometry  $f_{\mathbf{x}, \mathbf{y}}$ , and hence the set of its fixed points  $H(\mathbf{x}, \mathbf{y})$ , can be explicitly described (see, for example, [24, Example 4.6]). It follows from this explicit description that the submanifold  $H(\mathbf{x}, \mathbf{y})$  with the induced metric is again one of the three

model spaces with the same curvature as the ambient manifold  $M$  and having codimension one.

### 3.5 Evolution of the mirror isometries

Having classified the space  $M$ , we must now classify the set of drift vectorfields  $\mathbf{b}$  which permit *MMC* with *LPC*. This necessitates analysis of the evolution of the isometries  $F_s$  as  $s$  varies. As noted above, [30] proved that the set of isometries  $\mathcal{G}$  has the structure of a Lie group. The first objective is to prove that the curve of isometries  $(F_s : s \geq 0)$  is a  $C^1$  curve in this Lie group.

**Lemma 40** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. The curve  $s \mapsto F_s$  is a  $C^1$  curve in the Lie group  $\mathcal{G}$ .*

*Proof* Recall that any point in  $M$  has a neighbourhood, called a  $\sigma$ -neighbourhood, such that any point in this neighbourhood is in a normal coordinate ball of any other point in the same neighbourhood. We study continuity and continuous differentiability of  $(F_s : s \geq 0)$  at  $s = t$ . As we are investigating a local property, we work in two separate sets of normal coordinates; one set describing a  $\sigma$ -neighbourhood  $U$  around  $\mathbf{x}$  and the other set describing another  $\sigma$ -neighbourhood  $V$  around  $F_t(\mathbf{x})$  such that  $F_t(\overline{U}) \subset V$ .

The first step is to prove that  $s \mapsto F_s$  is continuous in  $\mathcal{G}$  at  $s = t < \tau$ . To show this, it suffices to show that any set of  $d + 1$  points  $\mathbf{x}_i \in M$ , all of which lie in a  $\sigma$ -neighbourhood and are linearly independent (i.e. do not belong in the same  $(d - 1)$ -dimensional geodesic hypersurface), produces continuous curves  $s \mapsto F_s(\mathbf{x}_i)$  in  $M$  [30]. We note here that we can obtain such a set of  $d + 1$  points in any dense subset of any open set in  $M$ . To show the continuity of these curves, we will use the continuity of the diffusion paths and the fact that, by Corollary 33,  $Y_s = F_s(X_s)$  when  $s < \tau$ .

Define the new distance

$$\overline{\text{dist}}(\mathbf{x}, \mathbf{y}) = \frac{\text{dist}(\mathbf{x}, \mathbf{y})}{1 + \text{dist}(\mathbf{x}, \mathbf{y})}$$

for  $\mathbf{x}, \mathbf{y} \in M$ . Note that  $\overline{\text{dist}}(\cdot, \cdot)$  is bounded and it defines a distance that produces the same topology on  $M$  as  $\text{dist}(\cdot, \cdot)$  does. Now, take any sequence  $\{s_n\}_{n \geq 1}$  with  $\lim_{n \rightarrow \infty} s_n = t$ . Then,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{E} [\overline{\text{dist}}(F_{s_n}(X_t), F_t(X_t)) \mathbb{I}(\tau > t)] \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{E} [\overline{\text{dist}}(F_{s_n}(X_t), F_{s_n}(X_{s_n})) \mathbb{I}(\tau > t)] \\ & \quad + \limsup_{n \rightarrow \infty} \mathbb{E} [\overline{\text{dist}}(F_{s_n}(X_{s_n}), F_t(X_t)) \mathbb{I}(\tau > t)] \\ & = \limsup_{n \rightarrow \infty} \mathbb{E} [\overline{\text{dist}}(X_t, X_{s_n}) \mathbb{I}(\tau > t)] + \limsup_{n \rightarrow \infty} \mathbb{E} [\overline{\text{dist}}(Y_{s_n}, Y_t) \mathbb{I}(\tau > t)] = 0. \end{aligned}$$



Here, the equality in the second step follows from the fact that  $F_{s_n}$  is an isometry,  $Y_s = F_s(X_s)$  when  $s < \tau$ , and the dominated convergence theorem. The last equality follows from the path continuity of  $X$  and  $Y$  and another application of the dominated convergence theorem. Thus,  $\overline{\text{dist}}(F_{s_n}(x), F_t(x))$  converges to zero in  $L^1$  with respect to the law of  $X_t$  restricted on  $\{\tau > t\}$ . Hence, we can extract a subsequence  $n_k$  such that  $F_{s_{n_k}}$  converges to  $F_t$  almost everywhere with respect to the same measure. As the law of  $X_t$  restricted on  $\{\tau > t\}$  has full support on  $H_t^-$ , therefore the set of  $\mathbf{x} \in H_t^-$  for which  $F_{s_{n_k}}(\mathbf{x}) \rightarrow F_t(\mathbf{x})$  is a dense subset of  $H_t^-$ . Hence, by the previous discussion,  $F_{s_{n_k}} \rightarrow F_t$  in  $\mathcal{G}$ . As the limit does not depend on the chosen subsequence  $n_k$ , we conclude that  $F_{s_n} \rightarrow F_t$  in  $\mathcal{G}$ , proving continuity of  $s \mapsto F_s$ .

It is necessary to address the question of right-continuity at  $t = 0$ . Take  $\mathbf{x} \in H_0^-$  and consider the case when  $t_n \downarrow 0$ . Take a sequence  $\mathbf{x}_n \rightarrow \mathbf{x}$  such that  $\mathbf{x}_n \in H_{t_n}^-$ . An argument following the treatment of the case  $s = 0$  in the proof of Lemma 35 shows that  $F_{t_n}(\mathbf{x}_n) \rightarrow F_0(\mathbf{x})$ . As  $F_{t_n}$  is an isometry for each  $n$ , we can deduce that  $F_{t_n}(\mathbf{x}) \rightarrow F_0(\mathbf{x})$ , thus proving right-continuity.

The next step is to prove differentiability at  $t > 0$ . With  $\sigma$ -neighbourhoods  $U, V$  of  $\mathbf{x}, F_t(\mathbf{x})$  as described above, let  $\tau_U = \inf\{s \geq t : X_s \notin U\}$ . Because the coupling is Markovian,  $\tau_U$  is a stopping time with respect to the filtration generated by the coupling process  $(X, Y)$ . Consider the stopped processes  $X_s^U = X_{s \wedge \tau_U}$  and  $Y_s^U = Y_{s \wedge \tau_U}$ . In a slight abuse of notation, we use the same notation  $X_s^U$  for the coordinate representation for this stopped process in  $U$ , and similarly for  $Y_s^U$ . Also we continue to write  $F_s$  for the coordinate representation of  $F_s : U \rightarrow V$ .

By Lemma 8 of [30] it suffices to prove differentiability at  $t$  of the continuous curve  $s \mapsto F_s(\mathbf{x})$  for  $\mathbf{x} \in H_t^-$  such that  $(\mathbf{x}, F_t(\mathbf{x})) \in \mathcal{M}(\mu_t)$ . Take  $U, V$  and normal coordinate systems for  $\mathbf{x}$  and  $F_t(\mathbf{x})$  as above. Using these coordinates, we may write the stochastic differential equation for  $X^U$  as

$$dX_s^{U,i} = \mathbf{b}^i(X_s^U) ds + \sum_{j=1}^d \sigma^{i,j}(X_s^U) dW_s^j$$

for some Brownian motion  $W$  in  $U$ . A similar expression holds for  $Y^U$  with  $\mathbf{b}_F^i$  and  $\sigma_F^{i,j}$  representing the corresponding quantities. General properties of diffusions [31, Chapter 11] yield the following expressions in coordinate form:

$$\begin{aligned} \mathbf{b}^i(\mathbf{x}) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{X_s^{U,i} - x^i}{s - t} \mid X_t^U = \mathbf{x} \right], \\ \sigma^{i,j}(\mathbf{x}) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{(X_s^{U,i} - x^i)(X_s^{U,j} - x^j)}{s - t} \mid X_t^U = \mathbf{x} \right], \\ \mathbf{b}_F^i(F_t(\mathbf{x})) &= \lim_{s \downarrow t} \mathbb{E} \left[ \frac{Y_s^{U,i} - F_t^i(\mathbf{x})}{s - t} \mid Y_t^U = F_t(\mathbf{x}) \right]. \end{aligned} \quad (63)$$

By Corollary 33,  $Y_s = F_s(X_s)$  when  $s < \tau$ . Thus, we can write

$$\begin{aligned} & \mathbb{E} \left[ \frac{F_s^i(X_s^U) - F_t^i(\mathbf{x})}{s-t} \mid Y_t^U = F_t(\mathbf{x}) \right] \\ &= \mathbb{E} \left[ \frac{F_s^i(X_s^U) - F_s^i(\mathbf{x})}{s-t} \mid X_t^U = \mathbf{x} \right] + \frac{F_s^i(\mathbf{x}) - F_t^i(\mathbf{x})}{s-t}. \end{aligned} \quad (64)$$

The third expression in (63) gives

$$\lim_{s \downarrow t} \mathbb{E} \left[ \frac{F_s^i(X_s^U) - F_t^i(\mathbf{x})}{s-t} \mid Y_t^U = F_t(\mathbf{x}) \right] = \mathbf{b}_F^i(F_t(\mathbf{x})).$$

As  $s \mapsto F_s$  is a continuous curve in  $\mathcal{G}$ , we may deduce by [30, Lemma 7] that the (space) derivatives of  $F_s$  are continuous in  $s$ . By a Taylor expansion of  $F_s$  in  $U$  based at  $\mathbf{x}$  and (63),

$$\begin{aligned} & \lim_{s \downarrow t} \mathbb{E} \left[ \frac{F_s^i(X_s^U) - F_s^i(\mathbf{x})}{s-t} \mid X_t^U = \mathbf{x} \right] \\ &= \lim_{s \downarrow t} \left( \sum_{j=1}^d \partial_j F_s^i(\mathbf{x}) \mathbb{E} \left[ \frac{X_s^{U,j} - x^j}{s-t} \mid X_t^U = \mathbf{x} \right] \right. \\ & \quad \left. + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_{j,k} F_s^i(\mathbf{x}) \mathbb{E} \left[ \frac{(X_s^{U,j} - x^j)(X_s^{U,k} - x^k)}{s-t} \mid X_t^U = \mathbf{x} \right] + o(1) \right) \\ &= \sum_{j=1}^d \partial_j F_t^i(\mathbf{x}) \mathbf{b}^j(\mathbf{x}) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_{j,k} F_t^i(\mathbf{x}) \sigma^{j,k}(\mathbf{x}). \end{aligned}$$

Thus, from (64), we deduce that the curve  $s \mapsto F_s(\mathbf{x})$  has a continuous right-derivative given by

$$\begin{aligned} \lim_{s \downarrow t} \frac{F_s^i(\mathbf{x}) - F_t^i(\mathbf{x})}{s-t} &= \mathbf{b}_F^i(F_t(\mathbf{x})) - \sum_{j=1}^d \partial_j F_t^i(\mathbf{x}) \mathbf{b}^j(\mathbf{x}) \\ & \quad - \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_{j,k} F_t^i(\mathbf{x}) \sigma^{j,k}(\mathbf{x}). \end{aligned} \quad (65)$$

This, together with [4, Theorem 1.3], implies uniformly continuous differentiability of  $s \mapsto F_s(\mathbf{x})$  at  $t > 0$ . Note that the Mean Value Theorem and right-continuity of the right hand side of (65) now gives us right-differentiability at  $t = 0$ . This proves the lemma.  $\square$

**Corollary 41** *All the partial derivatives with respect to  $\mathbf{x}$  of  $(s, \mathbf{x}) \mapsto F_s(\mathbf{x})$  are continuously differentiable in  $s$ . Furthermore,  $\frac{d}{ds}\big|_{s=t} F_s(\mathbf{x})$  is smooth in  $\mathbf{x}$ .*

*Proof* Using the argument of [30, Section 8], we can deduce the following representation in local coordinates  $(x^i)$ :

$$F_t(x^1, \dots, x^d) = \Psi(x^1, \dots, x^d, F_t(\mathbf{x}_0), \dots, F_t(\mathbf{x}_d)),$$

where  $\Psi$  is a smooth function and  $\mathbf{x}_0, \dots, \mathbf{x}_d$  are fixed points in  $M$ . The corollary follows from this representation and the previous lemma.  $\square$

The derivative vectorfield  $\kappa$  defined on  $M$  by

$$\kappa(\mathbf{x}) = \frac{d}{ds}\bigg|_{s=0} F_s(F_0(\mathbf{x}))$$

possesses a special significance. This is the *Killing vectorfield* corresponding to the  $C^1$  curve  $s \mapsto G_s$  in  $\mathcal{G}$  given by  $G_s(\mathbf{x}) = F_s(F_0(\mathbf{x}))$  for  $\mathbf{x} \in M$ . Vectorfields of this form correspond to the natural action of elements in the Lie algebra of  $\mathcal{G}$  on the manifold  $M$  (recall that  $F_0 \circ F_0$  is the identity map, and the Lie algebra of  $\mathcal{G}$  corresponds to the tangent space of  $\mathcal{G}$  at the identity). Killing vectorfields will play a crucial rôle in the following subsections.

### 3.6 Structure of the coupling

The processes  $X$  and  $Y$  can be constructed as projections  $X_t = \pi U_t$  and  $Y_t = \pi \tilde{U}_t$ , where  $U$  and  $\tilde{U}$  are solutions to Stratonovich stochastic differential equations which are defined on the orthonormal frame bundle  $\mathcal{O}(M)$  by

$$\begin{aligned} dU_t &= \sum_i H_i(U_t) \circ dW_t^i + \mathbf{B}(U_t) dt, \\ d\tilde{U}_t &= \sum_i H_i(\tilde{U}_t) \circ d\tilde{W}_t^i + \mathbf{B}(\tilde{U}_t) dt, \end{aligned} \quad (66)$$

for  $d$ -dimensional Euclidean Brownian motions  $W$  and  $\tilde{W}$  and the vectorfield and the lifted drift vectorfield  $\mathbf{B}$  given by (55).

Any isometry  $F$  on  $M$  has a natural lift to a smooth mapping  $\hat{F} : \mathcal{O}(M) \rightarrow \mathcal{O}(M)$ , given by

$$\hat{F}(\pi u, ue_1, \dots, ue_d) = (F(\pi u), dF(ue_1), \dots, dF(ue_d)). \quad (67)$$

The following lemma shows that  $\hat{F}$  respects the structure of horizontal vectorfields on  $\mathcal{O}(M)$ .

**Lemma 42** *Let  $F$  be an isometry on  $M$  and let  $\hat{F}$  be the lift to  $\mathcal{O}(M)$  as defined above. For  $1 \leq i \leq d$  and  $u \in \mathcal{O}(M)$ ,*

$$d \hat{F}(H_i(u)) = H_i(\hat{F}(u)).$$

*Proof* Let  $\gamma$  be the unit speed geodesic in  $M$  starting from  $\pi u$  in direction  $ue_i$ , defined on some interval  $[0, \varepsilon]$  for some  $\varepsilon > 0$ . For each  $1 \leq j \leq d$ , let  $u_t^j$  denote the parallel transport of  $ue_j$  along  $\gamma$ . Define the curve  $\gamma^u$  in  $\mathcal{O}(M)$  given by

$$\gamma^u(t) = (\gamma_t, u_t^1, \dots, u_t^d)$$

for  $t \in [0, \varepsilon]$ . As the covariant derivative commutes with the push-forward of vector fields by isometries [26, Proposition 5.6], for each  $1 \leq j \leq d$ ,  $d F(u_t^j)$  provides a parallel transport of  $d F(ue_j)$  along  $F \circ \gamma_t$ . Hence,

$$\gamma^{\hat{F}(u)}(t) = \hat{F} \circ \gamma^u(t) = \left( F \circ \gamma_t, d F(u_t^1), \dots, d F(u_t^d) \right).$$

Now  $(\gamma^u)'(0) = H_i(u)$ . Thus

$$\begin{aligned} d \hat{F}(H_i(u)) &= d \hat{F}((\gamma^u)'(0)) = (\hat{F} \circ \gamma^u)'(0) = (\gamma^{\hat{F}(u)})'(0) \\ &= H_i(\hat{F}(u)), \end{aligned}$$

proving the lemma.  $\square$

The stochastic differential equation (66) for  $U$  delivers a diffusion  $V$  on  $\mathcal{O}(M)$  given by

$$V_t = \hat{F}_t(U_t),$$

where  $F_t$  is the time-varying deterministic involutive isometry constructed in previous subsections. Note that this automatically implies  $Y_t = F_t(X_t) = \pi V_t$  on  $t < \tau$ . Thus,  $V$  lifts  $Y$  up to the orthonormal frame bundle  $\mathcal{O}(M)$ . We now derive the stochastic differential equation for  $V$ .

From [20, Equation (2.3)] it follows that

$$d V_t = \sum_i (d \hat{F}_t(H_i(U_t))) \circ d W_t^i + d \hat{F}_t(\mathbf{B}(U_t)) dt + \hat{\chi}_t(U_t) dt, \quad (68)$$

where

$$\hat{\chi}_t(u) = \left. \frac{d}{ds} \right|_{s=t} \hat{F}_s(u)$$

exists by Lemma 40 and Corollary 41.

Lemma 42 implies that

$$d\hat{F}_t(H_i(U_t)) = H_i(\hat{F}_t(U_t)) = H_i(V_t),$$

and

$$d\hat{F}_t(\mathbf{B}(U_t)) = \sum_i b_i(U_t) d\hat{F}_t(H_i(U_t)) = \sum_i b_i(\hat{F}_t(V_t)) H_i(V_t)$$

where we have used Lemma 42 and the fact that  $\hat{F}_t^2 = \text{Id}$  in the last step.

Thus, the stochastic differential equation for  $V$  takes the form

$$dV_t = \sum_i H_i(V_t) \circ dW_t^i + \sum_i b_i(\hat{F}_t(V_t)) H_i(V_t) dt + \hat{\chi}_t(\hat{F}_t(V_t)) dt. \quad (69)$$

Considering differentiation along the curve  $\gamma^u$  introduced in the proof of Lemma 42, it can be seen that

$$d\pi(H_i(u)) = ue_i.$$

Also, as  $F_t$  is an involutive isometry,

$$\begin{aligned} b_i(\hat{F}_t(V_t)) &= \langle \mathbf{b}(F_t(Y_t)), dF_t(V_t e_i) \rangle_{F_t(Y_t)} = \langle dF_t(\mathbf{b}(F_t(Y_t))), V_t e_i \rangle_{Y_t} \\ &= \langle F_{t*} \mathbf{b}(Y_t), V_t e_i \rangle_{Y_t}, \end{aligned}$$

where  $F_{t*} \mathbf{b}$  is the pushforward of the vectorfield  $\mathbf{b}$  on  $M$  by the isometry  $F_t$ .

Finally, writing

$$\chi_t(\mathbf{x}) = \left. \frac{d}{ds} \right|_{s=t} F_s(\mathbf{x})$$

for  $\mathbf{x} \in M$ , note that, for  $u \in \mathcal{O}(M)$  and a smooth function  $f : M \rightarrow \mathbb{R}$ ,

$$\begin{aligned} d\pi(\hat{\chi}_t(u))(f) &= \left. \frac{d}{ds} \right|_{s=t} (f \circ \pi \circ \hat{F}_s)(u) = \left. \frac{d}{ds} \right|_{s=t} f(F_s(\pi(u))) \\ &= \chi_t(\pi u)(f). \end{aligned}$$

Thus, writing

$$\kappa_t(\mathbf{x}) = \chi_t(F_t(\mathbf{x})) \quad (70)$$

for  $\mathbf{x} \in M$ , we obtain

$$d\pi(\hat{\chi}_t(\hat{F}_t(u))) = \kappa_t(\pi u).$$

Note that  $\kappa_t$  is the Killing vectorfield corresponding to the  $C^1$  curve of isometries  $(F_s \circ F_t : s \geq t)$ , as introduced at the end of Sect. 3.5.

Using the above relations, we can project down the stochastic differential equation (69) for  $V$  onto  $M$  as follows.

$$\begin{aligned} dY_t &= \sum_i d\pi(H_i(V_t)) \circ dW_t^i + \sum_i b_i(\hat{F}_t(V_t)) d\pi(H_i(V_t)) dt \\ &\quad + d\pi(\hat{\chi}_t(\hat{F}_t(V_t))) dt \\ &= \sum_i V_t e_i \circ dW_t^i + \sum_i \langle F_{t*} \mathbf{b}(Y_t), V_t e_i \rangle_{Y_t} V_t e_i dt + \kappa_t(Y_t) dt \\ &= \sum_i V_t e_i \circ dW_t^i + F_{t*} \mathbf{b}(Y_t) dt + \kappa_t(Y_t) dt. \end{aligned}$$

From the above expression, we see that the generator of  $Y$  at  $(t, \mathbf{x})$  is

$$L = \frac{1}{2} \Delta_M + F_{t*} \mathbf{b}(\mathbf{x}) + \kappa_t(\mathbf{x}).$$

Comparing this with (57), we deduce the following important relation:

**Theorem 43** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. For a Markovian maximal coupling  $(X, Y)$  to exist from starting points  $(\mathbf{x}_0, \mathbf{y}_0)$ , the following relation must hold:*

$$\mathbf{b}(\mathbf{x}) = F_{t*} \mathbf{b}(\mathbf{x}) + \kappa_t(\mathbf{x}) \quad (71)$$

for all  $\mathbf{x} \in M$  and  $t \geq 0$ , where  $(F_s : s \geq 0)$  is the  $C^1$  curve of isometries introduced in Lemma 35.

**Remark 44** If  $\mathbf{b} = 0$  in the above theorem, we get  $\kappa_t(\mathbf{x}) = 0$  for all  $\mathbf{x} \in M$  and all  $t \geq 0$ . In particular,  $\kappa_t(F_t(\mathbf{x})) = 0$ , which by (70) gives

$$\left. \frac{d}{ds} \right|_{s=t} F_s(\mathbf{x}) = 0$$

for all  $\mathbf{x} \in M$  and all  $t \geq 0$ . Thus,  $F_t \equiv F_0$  for all  $t \geq 0$ . As  $H_t$  is precisely the set of fixed points of  $F_t$ , we deduce that the mirror  $H_t$  does not depend on time  $t$ . This was also proved in [25, Proposition 4.2].

### 3.7 Classification of the drift

Finally it is possible to produce a complete characterization of the drift  $\mathbf{b}$  under *LPC*. Recall that  $M$  can only be a scaled version of one of the model spaces  $\mathbb{S}^d$ ,  $\mathbb{H}^d$  or  $\mathbb{R}^d$  corresponding to the curvature  $K$  being constant and equal to  $+1$ ,  $-1$ , or  $0$ .

For this section, special attention is paid to the Eq. (71) at time  $0$ . When the context makes it plain there is no ambiguity, we will write  $F$  for  $F_0$  and  $\kappa$  for  $\kappa_0$ .

Let  $\nabla$  represent the covariant derivative with respect to the Riemannian connection compatible with the metric  $g$ . We will need the following useful fact about Killing vectorfields [32, Prop. 27].

**Lemma 45** *If  $\kappa$  is a Killing vectorfield, then for any  $\mathbf{x} \in M$  and any  $u \in T_{\mathbf{x}}M$ ,*

$$\langle \nabla_u \kappa(\mathbf{x}), u \rangle = 0 \quad (72)$$

Isometries take geodesics to geodesics, so any Killing vectorfield is a *Jacobi field*, i.e. the variation field of a variation through geodesics. Thus, Killing vectorfields satisfy the Jacobi equation, as given by the following lemma [26, Theorem 10.2].

**Lemma 46** *Let  $\kappa$  be a Killing vectorfield. Then  $\kappa$  satisfies the Jacobi equation along any (unit speed) geodesic  $\gamma$ :*

$$\nabla_{\dot{\gamma}} \nabla_{\dot{\gamma}} \kappa + R(\kappa, \dot{\gamma}) \dot{\gamma} = 0. \quad (73)$$

Because of Theorem 38, we can confine attention to the case when  $M$  is of constant curvature  $K$ , in which case there is a simple representation for the curvature tensor  $R$  [26, Lemma 8.10]:

$$R(X, Y)Z = K(\langle Y, Z \rangle X - \langle X, Z \rangle Y). \quad (74)$$

We now define the *symmetric 2-form* associated with the drift vectorfield  $\mathbf{b}$ : for  $u, v \in T_{\mathbf{x}}M$ ,

$$S_{\mathbf{x}}(u, v) = \frac{1}{2} (\langle \nabla_u \mathbf{b}, v \rangle + \langle \nabla_v \mathbf{b}, u \rangle). \quad (75)$$

The following lemma describes this symmetric 2-form  $S_{\mathbf{x}}$  under *LPC*.

**Lemma 47** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Under LPC, there is a scalar  $\lambda$  such that, for all  $\mathbf{x} \in M$  and all  $u, v \in T_{\mathbf{x}}M$ ,*

$$S_{\mathbf{x}}(u, v) = \lambda \langle u, v \rangle.$$

*Proof* Recall that  $\mathbf{x}^*$  is the midpoint of a minimal geodesic connecting  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . Let  $\{e_1, \dots, e_d\}$  denote the canonical orthonormal frame of  $T_{\mathbf{x}^*}M$ . From previous discussions,  $F$  ‘inverts’ one geodesic through  $\mathbf{x}^*$  (the minimal geodesic joining  $\mathbf{x}_0$  and  $\mathbf{y}_0$ ) and keeps all geodesics orthogonal to this one fixed. Let  $\mathbf{n} \in T_{\mathbf{x}^*}M$  denote the direction of the inverted geodesic.

Now, consider any isometry  $G$  that satisfies

$$\mathbf{b}(\mathbf{x}) = G_* \mathbf{b}(\mathbf{x}) + \kappa(\mathbf{x}) \quad (76)$$

for some Killing vectorfield  $\kappa$ , for all  $\mathbf{x} \in M$ . Then, it follows that for any  $\mathbf{x} \in M$  and  $u, v \in T_{\mathbf{x}}M$ ,

$$\begin{aligned}\langle \nabla_u \mathbf{b}(\mathbf{x}), v \rangle &= \langle \nabla_u (G_* \mathbf{b})(\mathbf{x}), v \rangle + \langle \nabla_u \kappa(\mathbf{x}), v \rangle \\ &= \langle \nabla_{dG^{-1}(u)} \mathbf{b}(G^{-1}(\mathbf{x})), dG^{-1}(v) \rangle + \langle \nabla_u \kappa(\mathbf{x}), v \rangle\end{aligned}$$

which, along with Lemma 45, yields

$$S_{\mathbf{x}}(u, v) = S_{G^{-1}(\mathbf{x})}(dG^{-1}(u), dG^{-1}(v)). \quad (77)$$

In particular, Eq. (71) at time  $t = 0$  gives

$$S_{\mathbf{x}^*}(u, v) = S_{\mathbf{x}^*}(dF(u), dF(v)). \quad (78)$$

where (78) follows from (77) by noting that  $F$  fixes  $\mathbf{x}^*$  and  $F^{-1} = F$ . Let  $S(\mathbf{x}^*)$  denote the matrix

$$(S(\mathbf{x}^*))_{ij} = S_{\mathbf{x}^*}(e_i, e_j).$$

Using the description above of  $F$  as 'inverting' the geodesic with tangent vector  $\mathbf{n}$  at  $\mathbf{x}^*$ , and leaving orthogonal geodesics at  $\mathbf{x}^*$  fixed, (78) yields

$$S(\mathbf{x}^*) = (\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top)S(\mathbf{x}^*)(\mathbb{I} - 2\mathbf{n}\mathbf{n}^\top). \quad (79)$$

By *LPC*, we can choose  $d$  pairs of starting points  $\{(\mathbf{x}_i, \mathbf{y}_i) : \mathbf{x}_i \in \mathcal{B}(\mathbf{x}_0, r), \mathbf{y}_i \in \mathcal{B}(\mathbf{y}_0, r), 1 \leq i \leq d\}$  such that the directions of the inverted geodesics  $\mathbf{n}_i$  (for  $1 \leq i \leq d$ ) based at  $\mathbf{x}^*$  form  $d$  linearly independent vectors in  $T_{\mathbf{x}^*}M$  and  $\mathbf{n}_i$  is not orthogonal to  $\mathbf{n}_j$  for any  $i \neq j$ . Now, noting from Eq. (79) that  $\mathbf{n}_i$  are eigenvectors of  $S(\mathbf{x}^*)$ , we find

$$S(\mathbf{x}^*) = \lambda(\mathbf{x}^*)\mathbb{I} \quad (80)$$

for some scalar  $\lambda(\mathbf{x}^*)$ . In coordinate-free terms, this is the assertion of the lemma at point  $\mathbf{x}^*$ .

Now, we want to show that the assertion of the lemma holds at any  $\mathbf{x} \in M$ . Denote

$$\mathcal{Z} = \{G \in \mathcal{G} : G \text{ satisfies (76) for some Killing vectorfield } \kappa \text{ and all } \mathbf{x} \in M\}.$$

Recall that (77) holds for all  $G \in \mathcal{Z}$ . Thus, by (80), we get

$$S_{G^{-1}(\mathbf{x}^*)}(u, v) = \lambda(\mathbf{x}^*)\langle u, v \rangle$$

for all  $u, v \in T_{G^{-1}(\mathbf{x}^*)}M$ .

By continuity of the map

$$G \mapsto S_{G^{-1}(\mathbf{x}^*)}(dG^{-1}(u), dG^{-1}(v))$$

in the topology of isometries [30, Lemma 4], (77) holds for all  $G \in \overline{\mathcal{Z}}$ , where  $\overline{\mathcal{Z}}$  denotes the closed subgroup generated by  $\mathcal{Z}$ .



Now, from the developments in Sect. 3.3, observe that, under *LPC*, for any  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$ , there exists a unique involutive isometry  $f_{\mathbf{x}, \mathbf{y}}$  whose fixed point set is exactly the set  $H(\mathbf{x}, \mathbf{y})$ . These isometries satisfy (76) as this equation corresponds to (71) at time  $t = 0$  when the starting points of  $X$  and  $Y$  are taken to be  $\mathbf{x}$  and  $\mathbf{y}$  respectively. Furthermore, exactly along the lines of the proof of Lemma 36, we see that the orbit of  $\mathbf{x}^*$  under the closed subgroup of isometries generated by  $\{f_{\mathbf{x}, \mathbf{y}} : \mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r), \mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)\}$  is the whole of  $M$ . In particular, the orbit of  $\mathbf{x}^*$  under  $\overline{\mathcal{Z}}$  is  $M$ . Thus, for all  $\mathbf{x} \in M$ ,

$$S_{\mathbf{x}}(u, v) = \lambda(\mathbf{x}^*)\langle u, v \rangle$$

for all  $u, v \in T_{\mathbf{x}}M$ , proving the lemma.  $\square$

Now we describe the drift vectorfield along geodesics issuing from  $\mathbf{x}^*$ , the midpoint of a minimal geodesic joining  $\mathbf{x}_0$  and  $\mathbf{y}_0$ . In the following, we will denote the canonical orthonormal basis of  $T_{\mathbf{x}^*}M$  by  $\{e_1, \dots, e_d\}$ . Also, for any vector  $u \in T_{\mathbf{x}^*}M$  and any  $d \times d$  matrix  $T$ ,  $Tu$  will denote the vector obtained by matrix multiplication when we identify  $T_{\mathbf{x}^*}M$  with  $\mathbb{R}^d$ .

**Lemma 48** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. If the drift vectorfield  $\mathbf{b}$  permits MMC with LPC, then it must satisfy the following. Let  $\mathbf{x}^* \in M$  be the midpoint of a minimal geodesic connecting  $\mathbf{x}_0$  and  $\mathbf{y}_0$  and  $u, v \in T_{\mathbf{x}^*}M$  be unit vectors with  $u \perp v$ . Let  $\gamma$  represent the geodesic issuing from  $\mathbf{x}^*$  in direction  $u$  and let  $V_t$  represent the parallel transport of  $v$  along  $\gamma$ . Then the following holds.*

$$\langle \mathbf{b}(\gamma(t)), \dot{\gamma}_t \rangle = \lambda t + \langle \mathbf{b}(\mathbf{x}^*), u \rangle \quad (81)$$

where  $\lambda$  is as in Lemma 47, and

$$\langle \mathbf{b}(\gamma(t)), V_t \rangle = \begin{cases} \langle \mathbf{b}(\mathbf{x}^*), v \rangle \cos \sqrt{K}t + \langle Tu, v \rangle \frac{\sin \sqrt{K}t}{\sqrt{K}} & \text{if } K > 0, \\ \langle \mathbf{b}(\mathbf{x}^*), v \rangle + \langle Tu, v \rangle t & \text{if } K = 0, \\ \langle \mathbf{b}(\mathbf{x}^*), v \rangle \cosh \sqrt{-K}t + \langle Tu, v \rangle \frac{\sinh \sqrt{-K}t}{\sqrt{-K}} & \text{if } K < 0. \end{cases} \quad (82)$$

where the matrix  $T$  given by  $T_{ij} = \langle \nabla_{e_i} \mathbf{b}(\mathbf{x}^*), e_j \rangle - \lambda \langle e_i, e_j \rangle$  is a skew-symmetric matrix.

*Proof* To see (81), note that

$$\frac{d}{dt} \langle \mathbf{b}(\gamma(t)), \dot{\gamma}_t \rangle = \langle \nabla_{\dot{\gamma}_t} \mathbf{b}(\gamma(t)), \dot{\gamma}_t \rangle = S(\dot{\gamma}_t, \dot{\gamma}_t) = \lambda.$$

Take any  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  such that  $\mathbf{x}^* \in H(\mathbf{x}, \mathbf{y})$ . Since  $H(\mathbf{x}, \mathbf{y})$  is the fixed point set of the isometry  $f_{\mathbf{x}, \mathbf{y}}$ , it is therefore a totally geodesic submanifold of  $M$ . Let  $\kappa$  denote the Killing vectorfield for which (71) holds at time  $t = 0$  with

$F_0 = f_{\mathbf{x}, \mathbf{y}}$ . Take any unit speed geodesic  $\gamma$  passing through  $\mathbf{x}^*$  and lying in  $H(\mathbf{x}, \mathbf{y})$ . (Note that, if a geodesic lies in  $H(\mathbf{x}, \mathbf{y})$  for a short time, it should lie in  $H(\mathbf{x}, \mathbf{y})$  for all time. See, for example, the proof of Proposition 24 of [32], p. 145.)

Let  $(n_t : t \geq 0)$  be the parallel transport of the vector normal to the hypersurface  $H(\mathbf{x}, \mathbf{y})$  at  $\mathbf{x}^*$  along the geodesic  $\gamma$ . Note that, as  $H(\mathbf{x}, \mathbf{y})$  is totally geodesic, the second fundamental form vanishes identically on  $H(\mathbf{x}, \mathbf{y})$  [26, Exercise 8.4]. This fact implies that parallel transportation of a vector  $v \in T_{\mathbf{x}^*} H(\mathbf{x}, \mathbf{y})$  with respect to the induced metric on  $H(\mathbf{x}, \mathbf{y})$  agrees with parallel transportation of  $v$  in the ambient manifold  $M$  [26, Lemma 8.5]. Thus,  $n_t$  is precisely the direction that is reversed at  $\gamma(t)$  by  $f_{\mathbf{x}, \mathbf{y}}$ .

Equation (71) gives us

$$\langle \mathbf{b}(\gamma(t)), n_t \rangle = \frac{1}{2} \langle \kappa(\gamma(t)), n_t \rangle. \quad (83)$$

Differentiating the above twice with respect to  $t$  along the geodesic  $\gamma$ , and using the fact that  $\nabla_{\dot{\gamma}(t)} n_t = 0$  because  $n_t$  was defined using parallel transport along  $\gamma$ , we obtain

$$\langle D_t^2 \mathbf{b}(\gamma(t)), n_t \rangle = \frac{1}{2} \langle D_t^2 \kappa(\gamma(t)), n_t \rangle$$

(using  $D_t$  as shorthand for covariant differentiation  $\nabla_{\dot{\gamma}}$  along the geodesic  $\gamma$ ) which, along with (73) and (74), gives

$$\frac{d^2}{dt^2} \langle \mathbf{b}(\gamma(t)), n_t \rangle + \frac{K}{2} \langle \kappa(\gamma(t)), n_t \rangle = 0. \quad (84)$$

Consequently Eq. (83) shows that the function  $t \mapsto \langle \mathbf{b}(\gamma(t)), n_t \rangle$  satisfies the following differential equation

$$\frac{d^2}{dt^2} \langle \mathbf{b}(\gamma(t)), n_t \rangle + K \langle \mathbf{b}(\gamma(t)), n_t \rangle = 0. \quad (85)$$

For any geodesic  $\gamma$  passing through  $\mathbf{x}^*$ , not necessarily lying in  $H(\mathbf{x}, \mathbf{y})$ , and for any parallel vectorfield  $V_t$  along  $\gamma$  orthogonal to  $\dot{\gamma}_t$ , a similar technique uses (71), (73) and (74) to give us

$$\begin{aligned} & \frac{d^2}{dt^2} \langle \mathbf{b}(\gamma(t)), V_t \rangle + K \langle \mathbf{b}(\gamma(t)), V_t \rangle \\ &= \frac{d^2}{dt^2} \langle \mathbf{b}(f_{\mathbf{x}, \mathbf{y}} \circ \gamma(t)), d f_{\mathbf{x}, \mathbf{y}}(V_t) \rangle + K \langle \mathbf{b}(f_{\mathbf{x}, \mathbf{y}} \circ \gamma(t)), d f_{\mathbf{x}, \mathbf{y}}(V_t) \rangle. \end{aligned} \quad (86)$$

Now, following the lines of the proof of Lemma 37, we can iteratively compose the isometries in

$$\begin{aligned} \mathcal{S} &= \left\{ f_{\mathbf{x}, \mathbf{y}} \in \mathcal{G} : \mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r), \mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r), \text{dist}(\mathbf{x}, \mathbf{x}^*) = \text{dist}(\mathbf{y}, \mathbf{x}^*) \right. \\ &\quad \left. = \frac{1}{2} \text{dist}(\mathbf{x}, \mathbf{y}) \right\} \end{aligned}$$

to deduce that the closed subgroup of isometries  $\mathcal{G}^*$  generated by  $\mathcal{S}$  is the whole isotropy group of  $\mathbf{x}^*$  in  $\mathcal{G}$ . Further, from Step 1 and Step 2 in the proof of Lemma 37, it can be seen that for any pair of linearly independent unit vectors  $u, v \in T_{\mathbf{x}^*}M$ , there is a sequence of isometries  $\{F_k\}_{k \geq 1}$  such that for each  $k$ ,  $F_k$  is a composition of isometries in  $\mathcal{S}$ ,  $d F_k$  fixes vectors in  $T_{\mathbf{x}^*}M$  that are orthogonal to  $\{u, v\}$ , and  $d F_k(u) \rightarrow v$  as  $k \rightarrow \infty$ .

Take any geodesic  $\gamma$  issuing from  $\mathbf{x}^*$  and lying in  $H(\mathbf{x}, \mathbf{y})$  for some  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$ ,  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  and let  $n_t$  denote the parallel vectorfield along  $\gamma$  that is inverted by  $f_{\mathbf{x}, \mathbf{y}}$ . Let  $G \in \mathcal{G}$  be a composition of isometries in  $\mathcal{S}$  which fix  $\gamma$  and let  $v = d G(n_0)$ . Let  $V_t^v$  denote the parallel transport of  $v$  along  $\gamma$ . As  $G$  is an isometry, [26, Proposition 5.6 (b)] implies  $G_* n_t = V_t^v$ . Applying (86) at each composition corresponding to  $G$ , we get

$$\frac{d^2}{dt^2} \langle \mathbf{b}(\gamma(t)), n_t \rangle + K \langle \mathbf{b}(\gamma(t)), n_t \rangle = \frac{d^2}{dt^2} \langle \mathbf{b}(\gamma(t)), V_t^v \rangle + K \langle \mathbf{b}(\gamma(t)), V_t^v \rangle. \quad (87)$$

By (85), the left hand side of the above is zero. Thus, the right hand side should vanish too. Solving this gives (82) with  $V^v$  in place of  $V$  and the given matrix  $T$ .

Now, consider any parallel vectorfield  $V_t$  along  $\gamma$  which is orthogonal to  $\dot{\gamma}_t$ . By the discussion following the definition of  $\mathcal{S}$ , there exists a sequence of isometries  $\{F_k\}_{k \geq 1}$  such that each  $F_k$  is a composition of isometries in  $\mathcal{S}$ ,  $F_k$  fixes  $\gamma$ , and  $d F_k(n_0) \rightarrow V_0$  as  $k \rightarrow \infty$ . As  $F_k$  fixes  $\mathbf{x}^*$  for each  $k$ , by [30, p. 7], we can choose a subsequence  $k_l$  such that  $F_{k_l} \rightarrow F$  in  $\mathcal{G}$  as  $l \rightarrow \infty$ . Write  $V_t^{(k_l)} = F_{k_l*} n_t$ . By [30, Lemma 4], for each  $t \geq 0$ ,  $V_t^{(k_l)} \rightarrow d F(n_t)$  in  $T_{\gamma(t)}M$  as  $l \rightarrow \infty$ . In particular,  $d F(n_0) = V_0$ , and as  $F$  is an isometry fixing  $\gamma$ ,  $d F(n_t) = V_t$  for all  $t \geq 0$ . Thus, we have  $V_t^{(k_l)} \rightarrow V_t$  in  $T_{\gamma(t)}M$  for each  $t \geq 0$ . From the discussion in the previous paragraph, (82) holds with  $V^{(k_l)}$  in place of  $V$  for each  $l \geq 1$ . Taking  $l \rightarrow \infty$ , we obtain (82) for the vectorfield  $V$ .

Finally, take any pair of unit vectors  $u, v \in T_{\mathbf{x}^*}M$  satisfying  $u \perp v$ . Let  $\sigma$  be the geodesic issuing from  $\mathbf{x}^*$  such that  $\dot{\sigma}(0) = u$ . We can obtain a sequence of isometries  $\{G_k\}_{k \geq 1}$  such that each  $G_k$  is a composition of isometries in  $\mathcal{S}$  and  $d G_k(\dot{\gamma}(0)) \rightarrow u$  as  $k \rightarrow \infty$ . Write  $u_k = d G_k(\dot{\gamma}(0))$  and let  $\sigma_k$  be the geodesic issuing from  $\mathbf{x}^*$  in the direction  $u_k$ . Denote by  $V_t^{v,k}$  and  $V_t^v$  the parallel transport of  $v$  along  $\sigma_k$  and  $\sigma$  respectively. By the previous discussion, we know that (82) holds with  $V^{v,k}$  in place of  $V$  and  $\sigma_k$  in place of  $\gamma$  for each  $k \geq 1$ . Observe that for each fixed  $t \geq 0$ , both sides of (82) depend continuously on  $u$  and  $v$  (this observation for the left hand side follows from the fact that the solution to the geodesic and parallel transport equations depends continuously on the initial data). Thus, we can take  $k \rightarrow \infty$  to get (82) with  $V^v$  in place of  $V$  and  $\sigma$  in place of  $\gamma$ .

The fact that  $T$  is skew-symmetric follows from the observation that  $S_{\mathbf{x}^*}(e_i, e_j) = \lambda \langle e_i, e_j \rangle$  (by Lemma 47) and therefore

$$\langle \nabla_{e_i} \mathbf{b}(\mathbf{x}^*), e_j \rangle - \lambda \langle e_i, e_j \rangle = \frac{1}{2} (\langle \nabla_{e_i} \mathbf{b}(\mathbf{x}^*), e_j \rangle - \langle \nabla_{e_j} \mathbf{b}(\mathbf{x}^*), e_i \rangle).$$

□

Since  $M$  is a maximally symmetric space (by Theorem 38), the dimension of its set of Killing vectorfields is  $\frac{d(d+1)}{2}$ . Thus, for any vector  $w \in T_{\mathbf{x}^*}M$  and any skew-symmetric matrix  $T$ , there exists a unique Killing vectorfield  $\mathcal{K}$  with  $\mathcal{K}(\mathbf{x}^*) = w$  and  $\langle \nabla_{e_i} \mathcal{K}(\mathbf{x}^*), e_j \rangle = T_{ij}$ . Moreover, as every Killing vectorfield is a Jacobi field (i.e. satisfies (73)), it follows that  $\mathcal{K}$  satisfies the following equation analogous to (82), for unit vectors  $u, v \in T_{\mathbf{x}^*}M$  with  $u \perp v$ .

$$\langle \mathcal{K}(\gamma(t)), V_t \rangle = \begin{cases} \langle w, v \rangle \cos \sqrt{K}t + \langle Tu, v \rangle \frac{\sin \sqrt{K}t}{\sqrt{K}} & \text{if } K > 0, \\ \langle w, v \rangle + \langle Tu, v \rangle t & \text{if } K = 0, \\ \langle w, v \rangle \cosh \sqrt{-K}t + \langle Tu, v \rangle \frac{\sinh \sqrt{-K}t}{\sqrt{-K}} & \text{if } K < 0. \end{cases} \quad (88)$$

Thus, if we set  $\mathcal{K}_{\mathbf{x}^*}$  as the Killing vectorfield uniquely determined by  $w = \mathbf{b}(\mathbf{x}^*)$  and  $T_{ij} = \langle \nabla_{e_i} \mathbf{b}(\mathbf{x}^*), e_j \rangle - \lambda \langle e_i, e_j \rangle$ , we see from Lemmas 47 and 48 that the vectorfield  $\mathbf{b}$  can be written as

$$\mathbf{b} = \mathcal{D}_{\mathbf{x}^*}^\lambda + \mathcal{K}_{\mathbf{x}^*} \quad (89)$$

where  $\mathcal{D}_{\mathbf{x}^*}^\lambda$  is the *dilation vectorfield about  $\mathbf{x}^*$*  with dilation coefficient  $\lambda$  defined as

$$\mathcal{D}_{\mathbf{x}^*}^\lambda(\gamma(t)) = \lambda t \dot{\gamma}(t) \quad (90)$$

for any geodesic  $\gamma$  issuing from  $\mathbf{x}^*$ . Now, we claim that dilation vectorfields do not arise in the case of non-zero-curvature.

**Lemma 49**  $K \neq 0$  implies  $\lambda = 0$ .

*Proof* Under LPC, the description of  $\mathbf{b}$  given in Lemma 48 holds for  $\mathbf{x}^*$  replaced by  $\hat{x} \in \mathcal{B}(\mathbf{x}^*, \rho)$  for some  $\rho > 0$ . Take any two points  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{B}(\mathbf{x}^*, \rho)$  with  $\mathbf{x}_1 \neq \mathbf{x}_2$ . Lemmas 47 and 48, applied at  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , show that  $\mathbf{b}$  satisfies

$$\mathbf{b} = \mathcal{D}_1^\lambda + \mathcal{K}_1 = \mathcal{D}_2^\lambda + \mathcal{K}_2 \quad (91)$$

where  $\mathcal{K}_1$  and  $\mathcal{K}_2$  are Killing vectorfields and  $\mathcal{D}_1^\lambda$  and  $\mathcal{D}_2^\lambda$  are dilation vectorfields with the same coefficient  $\lambda$  about  $\mathbf{x}_1$  and  $\mathbf{x}_2$  respectively.

Denote by  $\sigma$  the geodesic issuing from  $\mathbf{x}_2$  and passing through  $\mathbf{x}_1$ , and set  $\gamma$  to be a geodesic issuing from  $\mathbf{x}_2$  in a direction orthogonal to  $\sigma$ . Locate  $\mathbf{z} = \gamma(\text{dist}(\mathbf{x}_1, \mathbf{x}_2))$ . Taking  $\rho$  sufficiently small, we can ensure that  $\gamma$  restricted to  $[0, \text{dist}(\mathbf{x}_1, \mathbf{x}_2)]$  is a

minimal geodesic from  $\mathbf{x}_2$  to  $\mathbf{z}$ . Finally, denote the geodesic issuing from  $\mathbf{x}_1$  and passing through  $\mathbf{z}$  by  $\eta$ . Consider the geodesic triangle  $\Delta$  formed by  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  and  $\mathbf{z}$ . Thus, the sides of  $\Delta$  are formed by the geodesics  $\sigma$ ,  $\gamma$  and  $\eta$ .

Now, recall that the curvature  $K$  can also be interpreted in terms of the rate at which geodesics diverge when they issue from a point in different directions. Thus [28, Proposition 2.6] we see that if  $\mathbf{x}_1$  is taken sufficiently close to  $\mathbf{x}_2$ , then

$$\begin{aligned}\text{dist}(\mathbf{x}_1, \mathbf{z}) &< \sqrt{2} \text{dist}(\mathbf{x}_1, \mathbf{x}_2) \text{ if } K > 0, \\ \text{dist}(\mathbf{x}_1, \mathbf{z}) &> \sqrt{2} \text{dist}(\mathbf{x}_1, \mathbf{x}_2) \text{ if } K < 0.\end{aligned}\quad (92)$$

Applying the triangle version of the Toponogov comparison theorem [32, Theorem 79, p. 339], we see that the interior angle  $\theta$  formed at the vertex  $\mathbf{z}$  of  $\Delta$  satisfies  $\theta \geq \pi/4$  if  $K > 0$  and  $\theta \leq \pi/4$  if  $K < 0$ . But (90) implies

$$\begin{aligned}\langle D_1^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle &= \langle D_1^\lambda(\mathbf{z}), \dot{\eta}(\text{dist}(\mathbf{x}_1, \mathbf{z})) \rangle \cos \theta \\ &= \lambda \text{dist}(\mathbf{x}_1, \mathbf{z}) \cos \theta.\end{aligned}$$

Thus, if  $\lambda > 0$ , we get

$$\begin{aligned}\langle D_1^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle &< \lambda \text{dist}(\mathbf{x}_1, \mathbf{x}_2) \text{ if } K > 0, \\ \langle D_1^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle &> \lambda \text{dist}(\mathbf{x}_1, \mathbf{x}_2) \text{ if } K < 0.\end{aligned}\quad (93)$$

and the inequalities are reversed if  $\lambda < 0$ .

From (91)

$$\begin{aligned}\langle D_2^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle &= \langle D_1^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle \\ &\quad + \langle (\mathcal{K}_1 - \mathcal{K}_2)(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle.\end{aligned}\quad (94)$$

Lemma 45 implies that the inner product of a Killing vectorfield with the velocity vector of a geodesic is conserved along the geodesic, yielding

$$\langle (\mathcal{K}_1 - \mathcal{K}_2)(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle = \langle (\mathcal{K}_1 - \mathcal{K}_2)(\mathbf{x}_2), \dot{\gamma}(0) \rangle.$$

From (90) it follows that  $D_2^\lambda(\mathbf{x}_2) = 0$  and also

$$\langle D_1^\lambda(\mathbf{x}_2), \dot{\gamma}(0) \rangle = \lambda \text{dist}(\mathbf{x}_1, \mathbf{x}_2) \langle \dot{\sigma}(0), \dot{\gamma}(0) \rangle = 0.$$

Combining this with (91),

$$\langle (\mathcal{K}_1 - \mathcal{K}_2)(\mathbf{x}_2), \dot{\gamma}(0) \rangle = 0.$$

Thus, (94) gives us

$$\langle D_2^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle = \langle D_1^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle.$$

By (90),  $\langle D_2^\lambda(\mathbf{z}), \dot{\gamma}(\text{dist}(\mathbf{x}_1, \mathbf{x}_2)) \rangle = \lambda \text{dist}(\mathbf{x}_1, \mathbf{x}_2)$ . Together with (93), this forces  $\lambda = 0$  if the curvature is non-zero, hence proving the lemma.  $\square$

*Note* When  $K > 0$ , observe that

$$\langle \mathbf{b}(\gamma(0)), \dot{\gamma}_0 \rangle = \left\langle \mathbf{b}(\gamma(2\pi/\sqrt{K})), \dot{\gamma}_{2\pi/\sqrt{K}} \right\rangle$$

yields  $\lambda = 0$ . But the above proof works for both positive and negative curvatures, and is in some sense, the real geometric reason why the dilation part of the vectorfield  $\mathbf{b}$  vanishes for non-zero curvature.

Finally we can state and prove the main theorem of this section.

**Theorem 50** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. The drift vectorfield  $\mathbf{b}$  permits MMC with LPC if and only if both of the following hold:*

- (i) *The underlying Riemannian manifold  $M$  is one of the three model spaces  $\mathbb{S}^d$  ( $K > 0$ ),  $\mathbb{R}^d$  ( $K = 0$ ) or  $\mathbb{H}^d$  ( $K < 0$ ), in the sense that the diffusion must be expressible as Riemannian Brownian motion plus drift vectorfield  $\mathbf{b}$  for such an  $M$ .*
- (ii) *For  $K \neq 0$ , the drift  $\mathbf{b}$  must and can be any Killing vectorfield  $\mathcal{K}$  on  $M$ . For  $K = 0$ , the drift  $\mathbf{b}$  must and can be described in Euclidean coordinates by  $\mathbf{b}(\mathbf{x}) = \lambda \mathbf{x} + T\mathbf{x} + \mathbf{c}$  for any scalar  $\lambda$ , any skew-symmetric matrix  $T$  and any vector  $\mathbf{c}$ , where  $\mathbf{x} \mapsto \lambda \mathbf{x}$  is a dilation vectorfield about the origin and  $\mathbf{x} \mapsto T\mathbf{x} + \mathbf{c}$  is a Killing vectorfield.*

*Proof* The classification of the space  $M$  is essentially the content of Theorem 38. Lemmas 48 and 49 show that if LPC holds then the drift vectorfield  $\mathbf{b}$  has to be of the form described in the theorem. For the case  $K = 0$ , Sect. 2 shows the existence of a Markovian maximal coupling with any pair of starting points  $\mathbf{x} \in \mathcal{B}(\mathbf{x}_0, r)$  and  $\mathbf{y} \in \mathcal{B}(\mathbf{y}_0, r)$  and fully describes the coupling.

To show existence and to describe the coupling for  $K \neq 0$ , recall that any Killing vectorfield  $\mathcal{K}$  generates a one-parameter subgroup of isometries starting from the identity, say  $(\Upsilon_t : t \in \mathbb{R})$ . Let  $Z$  denote a Brownian motion on  $M$ , and consider the law of

$$X_t = \Upsilon_t(Z_t).$$

Consider the lift  $U$  of the Brownian motion  $Z$  onto the orthonormal frame bundle  $\mathcal{O}(M)$ . Recall that the Stratonovich stochastic differential equation for this lifted process is given by

$$dU_t = \sum_i H_i(U_t) \circ dW_t^i \quad (95)$$

where  $W = (W^1, \dots, W^d)$  is a  $d$ -dimensional Euclidean Brownian motion. The process  $Z$  is recovered from  $U$  by  $Z = \pi(U)$ . Recall that the lift of an isometry  $F$  on  $M$  to  $\hat{F}$  on  $\mathcal{O}(M)$  is given by (67). Defining the process  $V$  on  $\mathcal{O}(M)$  by

$$V_t = \hat{\Upsilon}_t(U_t) \quad (96)$$

the arguments used to derive (69) also show that the Stratonovich stochastic differential equation for  $V$  is given by

$$dV_t = \sum_i H_i(V_t) \circ dW_t^i + \hat{\mathcal{K}}_t(\hat{\Upsilon}_t^{-1}(V_t)) dt \quad (97)$$

where

$$\hat{\mathcal{K}}_t(u) = \left. \frac{d}{ds} \right|_{s=t} \hat{\Upsilon}_s(u)$$

for  $u \in \mathcal{O}(M)$ . Note that, for any  $\mathbf{x} \in M$ ,

$$\left. \frac{d}{ds} \right|_{s=t} \Upsilon_s(\mathbf{x}) = \mathcal{K}(\Upsilon_t(\mathbf{x})).$$

Using this, and the fact that  $\pi(V_t) = X_t$ , we see that

$$\begin{aligned} dX_t &= \sum_i (d\pi(H_i(V_t))) \circ dW_t^i + (d\pi(\hat{\mathcal{K}}_t(\hat{\Upsilon}_t^{-1}(V_t)))) dt \\ &= \sum_i V_t e_i \circ dW_t^i + \mathcal{K}(X_t) dt \end{aligned}$$

which demonstrates that  $X$  is a Riemannian Brownian motion with drift vectorfield given by the Killing vectorfield  $\mathcal{K}$ .

As discussed in [25, Example 6.1] and references therein, if  $M$  is  $\mathbb{S}^d$  or  $\mathbb{H}^d$  then there exists a Markovian maximal coupling  $(Z, \tilde{Z})$  of Brownian motions starting from any two distinct points on  $M$ . Consider a diffusion representable as Riemannian Brownian motion with drift given by any Killing vectorfield  $\mathcal{K}$  on such a manifold  $M$ . Thus Lemma 4 implies that a Markovian maximal coupling for this diffusion exists between any pair of starting points, and can be constructed by

$$((\Upsilon_t(Z_t), \Upsilon_t(\tilde{Z}_t)) : t \geq 0)$$

where  $(\Upsilon_t : t \in \mathbb{R})$  is the one-parameter subgroup of isometries starting from the identity which is generated by the Killing vectorfield  $\mathcal{K}$ . This proves the theorem.  $\square$

**Corollary 51** *Under the hypothesis of part (ii) of Theorem 50, let  $(\Upsilon_t : t \in \mathbb{R})$  denote the one-parameter subgroup of isometries corresponding to the Killing vectorfield  $\mathcal{K}$ . Then for  $t \geq 0$ , the mirror  $H_t$  and the corresponding reflection isometries  $F_t$  satisfy  $H_t = \Upsilon_t(H_0)$  and  $F_t = \Upsilon_t \circ F_0 \circ \Upsilon_t^{-1}$ .*

*Proof* Let  $Z, \tilde{Z}$  be maximally coupled Brownian motions on  $M$ . For any  $t \geq 0$ , by Remark 44,  $H_0 = H(Z_t, \tilde{Z}_t)$  almost surely. By Theorem 32,  $H_t = H(\Upsilon(Z_t), \Upsilon(\tilde{Z}_t))$  almost surely. From this,  $H_t = \Upsilon_t(H_0)$  easily follows. Further, as  $F_t$  and  $\Upsilon_t \circ F_0 \circ \Upsilon_t^{-1}$

have the same set of fixed points, namely  $H_t$ , and neither of them is the identity, therefore  $F_t = \Upsilon_t \circ F_0 \circ \Upsilon_t^{-1}$  follows from uniqueness of isometry with fixed point set  $H_t$ .  $\square$

In the following theorem, we characterise the class of drifts  $\mathbf{b}$  and starting points  $\mathbf{x}_0, \mathbf{y}_0$  for which the interface  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  does not depend on time  $t$ .

**Theorem 52** *Suppose that the standing assumptions of diffusion-geodesic completeness and stochastic completeness both hold. Suppose the drift vectorfield  $\mathbf{b}$  permits MMC with LPC. Let  $I(\mathbf{x}_0, \mathbf{y}_0, t)$  denote the interface for the MMC  $(X, Y)$  of diffusions  $X$  and  $Y$  starting from  $\mathbf{x}_0$  and  $\mathbf{y}_0$  respectively. Then  $I(\mathbf{x}_0, \mathbf{y}_0, t) = I(\mathbf{x}_0, \mathbf{y}_0, 0)$  for all  $t \geq 0$  if and only if one of the following holds:*

- (i)  $K = 0$ ,  $\mathbf{b}(\mathbf{x}) = \lambda \mathbf{x} + T\mathbf{x} + \mathbf{c}$  for some scalar  $\lambda$ , skew-symmetric matrix  $T$  and vector  $\mathbf{c}$ , and  $\mathbf{x}_0, \mathbf{y}_0, \lambda, T, \mathbf{c}$  satisfy  $T(\mathbf{x}_0 - \mathbf{y}_0) = 0$  and  $(\mathbf{x}_0 - \mathbf{y}_0)^\top (\lambda(\mathbf{x}_0 + \mathbf{y}_0) + 2\mathbf{c}) = 0$ .
- (ii)  $K \neq 0$  and  $\mathbf{b}$  is a Killing vectorfield  $\mathcal{K}$  on  $M$  which satisfies the following: if  $\mathbf{x}^*$  is the midpoint of a minimal geodesic joining  $\mathbf{x}_0$  and  $\mathbf{y}_0$  and  $\mathbf{n}$  is the vector normal to the hypersurface  $H(\mathbf{x}_0, \mathbf{y}_0)$  at  $\mathbf{x}^*$ , then  $\langle \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = 0$  and  $\nabla_n \mathcal{K}(\mathbf{x}^*) = 0$ .

*Proof* When  $K = 0$ , we observe from (42) that  $\mathbf{n}(t) = \mathbf{n}(0)$  for all  $t \geq 0$  if and only if  $T(\mathbf{x}_0 - \mathbf{y}_0) = 0$ . Using this in (43), we get for  $\lambda \neq 0$ ,

$$l(t) = l(0)e^{\lambda t} + \frac{\mathbf{n}(0)^\top \mathbf{c}}{\lambda}(e^{\lambda t} - 1) = e^{\lambda t} \left( l(0) + \frac{\mathbf{n}(0)^\top \mathbf{c}}{\lambda} \right) - \frac{\mathbf{n}(0)^\top \mathbf{c}}{\lambda}.$$

Thus  $l(t) = l(0)$  for all  $t \geq 0$  if and only if  $l(0) + \frac{\mathbf{n}(0)^\top \mathbf{c}}{\lambda} = 0$ . Substituting  $l(0) = \frac{|\mathbf{x}_0|^2 - |\mathbf{y}_0|^2}{2|\mathbf{x}_0 - \mathbf{y}_0|}$  and  $\mathbf{n}(0) = \frac{\mathbf{x}_0 - \mathbf{y}_0}{|\mathbf{x}_0 - \mathbf{y}_0|}$  in this equation, we get  $(\mathbf{x}_0 - \mathbf{y}_0)^\top (\lambda(\mathbf{x}_0 + \mathbf{y}_0) + 2\mathbf{c}) = 0$ .

When  $\lambda = 0$ , we get  $l(t) = l(0) + t(\mathbf{n}(0)^\top \mathbf{c})$ . Thus  $l(t) = l(0)$  for all  $t \geq 0$  if and only if  $(\mathbf{x}_0 - \mathbf{y}_0)^\top \mathbf{c} = 0$ .

Now, suppose  $K \neq 0$  and  $\mathbf{b}$  is the Killing vectorfield  $\mathcal{K}$  on  $M$ . As there is at most one isometry whose fixed point set is  $H(\mathbf{x}_0, \mathbf{y}_0)$ , we deduce that  $I(\mathbf{x}_0, \mathbf{y}_0, t) = I(\mathbf{x}_0, \mathbf{y}_0, 0)$  for all  $t \geq 0$  if and only if  $F_t = F$  for all  $t \geq 0$ .

Suppose  $F_t = F$  for all  $t \geq 0$ . Then by (71),  $\mathcal{K}(\mathbf{x}) = F_* \mathcal{K}(\mathbf{x})$  for all  $\mathbf{x} \in M$ . In particular,  $\langle \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = \langle F_* \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle$ . But, as  $F$  is an involutive isometry,  $\langle F_* \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = \langle \mathcal{K}(\mathbf{x}^*), F_* \mathbf{n} \rangle = \langle \mathcal{K}(\mathbf{x}^*), -\mathbf{n} \rangle$  from which we get  $\langle \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = 0$ . Now, observe that as  $\mathcal{K}$  is a Killing vectorfield, therefore by Lemma 45,  $\langle \nabla_n \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = 0$ . If  $u \in T_{\mathbf{x}^*} M$  is orthogonal to  $\mathbf{n}$ , then

$$\begin{aligned} \langle \nabla_n \mathcal{K}(\mathbf{x}^*), u \rangle &= \langle \nabla_n F_* \mathcal{K}(\mathbf{x}^*), u \rangle = \langle F_* \nabla_{-\mathbf{n}} \mathcal{K}(\mathbf{x}^*), u \rangle = \langle \nabla_{-\mathbf{n}} \mathcal{K}(\mathbf{x}^*), F_* u \rangle \\ &= \langle -\nabla_n \mathcal{K}(\mathbf{x}^*), u \rangle \end{aligned}$$

which gives  $\langle \nabla_n \mathcal{K}(\mathbf{x}^*), u \rangle = 0$ . Hence,  $\langle \nabla_n \mathcal{K}(\mathbf{x}^*), u \rangle = 0$  for all  $u \in T_{\mathbf{x}^*} M$ , and therefore,  $\nabla_n \mathcal{K}(\mathbf{x}^*) = 0$ .

Conversely, suppose  $\langle \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = 0$  and  $\nabla_n \mathcal{K}(\mathbf{x}^*) = 0$  holds. Let  $\gamma$  be any geodesic issuing from  $\mathbf{x}^*$  and lying in  $H(\mathbf{x}_0, \mathbf{y}_0)$  and let  $n_t$  denote the parallel transport of  $\mathbf{n}$  along  $\gamma$ . As  $\langle \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = 0$  and  $\langle \nabla_{\dot{\gamma}(0)} \mathcal{K}(\mathbf{x}^*), \mathbf{n} \rangle = -\langle \nabla_n \mathcal{K}(\mathbf{x}^*), \dot{\gamma}(0) \rangle = 0$ , using



the representation (88) for  $\mathcal{K}$ , we see that  $\langle \mathcal{K}(\gamma(t), n_t) \rangle = 0$  and hence,  $\mathcal{K}(\gamma(t)) \in T_{\gamma(t)}H(\mathbf{x}_0, \mathbf{y}_0)$  for all  $t \geq 0$ . As the submanifold  $H(\mathbf{x}_0, \mathbf{y}_0)$  is a geodesic space, we conclude that  $\mathcal{K}$  restricted to  $H(\mathbf{x}_0, \mathbf{y}_0)$  is a vectorfield tangent to this submanifold. Thus, if  $\Upsilon_t$  denotes the flow of isometries generated by  $\mathcal{K}$ , then for each  $\mathbf{z}_0 \in H(\mathbf{x}_0, \mathbf{y}_0)$ ,  $\Upsilon_t(\mathbf{z}_0)$  lies in  $H(\mathbf{x}_0, \mathbf{y}_0)$  at least for a short time. As  $\Upsilon_t$  is a global flow (because  $M$  is complete), a routine compactness argument implies that  $\Upsilon_t(\mathbf{z}_0) \in H(\mathbf{x}_0, \mathbf{y}_0)$  for all  $t \geq 0$ . Thus, by Corollary 51,  $H_t \subseteq H(\mathbf{x}_0, \mathbf{y}_0)$ , and hence  $F_t = F$ , for all  $t \geq 0$ .  $\square$

## 4 Conclusion

In this paper we have shown that Markovian maximal couplings of regular elliptic diffusions with smooth coefficients (and satisfying diffusion-geodesic completeness and stochastic completeness) have to be reflection couplings tied to involutive isometries of the corresponding Riemannian structure on state space; moreover as soon as the existence of a Markovian maximal coupling is stable (in the sense of *LPC*) then a rigidity result requires the Riemannian structure to be Euclidean, hyperspherical, or hyperbolic, and the space must be simply connected. In such cases the drift must also be of a very simple form, corresponding to a rotation with possibly (but only in the Euclidean case) a dilation component.

Thus Markovian maximal couplings of elliptic diffusions are rare, and their existence enforces severe geometric constraints.

It is natural to ask whether the assumptions of diffusion-geodesic completeness and stochastic completeness are required. It seems likely that they are not required, but (this paper already being long) we save this question for another occasion.

The scarcity of Markovian maximal couplings places a natural premium on questions of efficiency of Markovian coupling, as discussed for example in [6], for the case of reflecting Brownian motion in compact regions. One could ask, for example, when it is possible to construct Markovian couplings  $(X, Y)$  which are optimal in the sense that the tail probability of the coupling time  $\mathbb{P}[\tau > t]$  is minimized for all  $t$  amongst Markovian couplings if not amongst all possible couplings. (Note that this notion of optimality differs from the optimality discussed in [8], which is defined relative to a specified Wasserstein metric.) Little is known as yet about such couplings, though [22] exhibits a coupling of two copies of scalar Brownian motion and local time which is Markovian, non-maximal, but optimal amongst all Markovian couplings. The question of whether similar geometric rigidity results for existence of such optimal Markovian couplings remains entirely open, and its answer would be of great interest.

We expect that in fact such optimal Markovian couplings are also rare. Further refinements are possible (for example, one could consider the existence of Markovian couplings which minimize the Laplace transform  $\mathbb{E}[\exp(-u\tau)]$  for some or all values of  $u > 0$ ); however the probable rarity of such couplings would focus attention on developing the notions of efficiency from [6] to apply to non-compact regions. In particular there is a natural question concerning criteria for existence of efficient Markovian couplings, where “efficient” here means, the rate of decay of  $\mathbb{P}[\tau > t]$  with  $t$  for the Markovian coupling is comparable to that of the total variation distance

$\|\mu_{1,t} - \mu_{2,t}\|_{TV}$  between the one-point distributions  $\mu_{1,t}$  and  $\mu_{2,t}$  (the distributions of  $X_t$  and  $Y_t$  respectively).

Two other natural extensions of these results are:

1. extension of the notion of Markovian maximal coupling to the hypoelliptic case (in which case in fact the very existence of Markovian couplings is moot: but see the positive results of [21, 23]);
2. examination of the extent to which the ideas of this paper carry over to Markov processes which are not skip-free (and here a natural first step would be to consider the case of couplings of Lévy processes, though a potentially significant result in the random walk case is to be found in [35]).

We hope to consider many of these questions in future work.

**Acknowledgments** We wish to thank an anonymous referee whose very careful reading of the manuscript and detailed comments greatly improved the article.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Benjamini, I., Burdzy, K., Chen, Z.-Q.: Shy couplings. *Probab. Theory Related Fields* **137**(3–4), 345–377 (2007)
2. Bramson, M., Burdzy, K., Kendall, W.: Shy couplings, CAT(0) spaces, and the Lion and Man. *Ann. Probab.* **41**(2), 744–784 (2013)
3. Bramson, M., Burdzy, K., Kendall, W.S.: Rubber bands, pursuit games and shy couplings. In: *Proceedings of the London Mathematical Society*, p. 48 (2014)
4. Bruckner, A.M.: *Differentiation of Real Functions. Lecture Notes in Mathematics*. American Mathematical Soc, Providence (1978)
5. Burago, D., Burago, Y., Ivanov, S.: *A Course in Metric Geometry, Graduate Studies in Mathematics*, vol. 33. American Mathematical Society, Providence (2001)
6. Burdzy, K., Kendall, W.S.: Efficient Markovian couplings: examples and counterexamples. *Ann. Appl. Probab.* **10**(2), 362–409 (2000)
7. Chavel, I.: *Riemannian Geometry: A Modern Introduction*. Number 108 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1995)
8. Chen, M.-F.: *From Markov Chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific Publishing Co., Inc, River Edge (2004)
9. Connor, S.B.: *Coupling: Cutoffs, CFTP and Tameness*. Phd thesis, University of Warwick (2007)
10. Cranston, M.: Gradient estimates on manifolds using coupling. *J. Funct. Anal.* **99**(1), 110–124 (1991)
11. Dynkin, E.B.: Markov Processes. Number 121/122 in *Die Grundlehren der Mathematischen Wissenschaften*. Springer, New York (1965)
12. Elworthy, K.D.: *Stochastic Differential Equations on Manifolds. LMS Lecture Note Series*. CUP (1982)
13. Gallier, J.: Spectral theorems in Euclidean and Hermitian spaces. In: *Geometric Methods and Applications*, pp. 343–365. Springer, New York (2011)
14. Gallot, S., Hulin, D., Lafontaine, J.: *Riemannian Geometry*, 3rd edn. Universitext, Springer, Berlin (2004)
15. Goldstein, S.: Maximal coupling. *Probab. Theory Related Fields* **46**(2), 193–204 (1979)
16. Griffiths, D.: A maximal coupling for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31**, 95–106 (1974/1975)
17. Hsu, E.P.: *Stochastic Analysis on Manifolds, Graduate Studies in Mathematics*, vol. 38. American Mathematical Society, Providence (2002)

18. Hsu, E.P., Sturm, K.-T.: Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.* **1**(1), 93–104 (2013)
19. Kendall, W.S.: Nonnegative Ricci curvature and the Brownian coupling property. *Stoch. Stoch. Rep.* **19**(1–2), 111–129 (1986)
20. Kendall, W.S.: Stochastic differential geometry: an introduction. *Acta Applicandae Mathematica* **9**(1–2), 29–60 (1987)
21. Kendall, W.S.: Coupling all the Lévy stochastic areas of multidimensional Brownian motion. *Ann. Probab.* **35**(3), 935–953 (2007)
22. Kendall, W.S.: Coupling, local times, immersions. *Bernoulli* **21**(2), 1014–1046 (2015)
23. Kendall, W.S., Price, C.J.: Coupling iterated Kolmogorov diffusions. *Electron. J. Probab.* **9**(Paper 13), 382–410 (2004)
24. Kuwada, K.: On uniqueness of maximal coupling for diffusion processes with a reflection. *J. Theor. Probab.* **20**(4), 935–957 (2007)
25. Kuwada, K.: Characterization of maximal Markovian couplings for diffusion processes. *Electron. J. Probab.* **14**(25), 633–662 (2009)
26. Lee, J.M.: *Riemannian Manifolds: An Introduction to Curvature*, vol. 176. Springer, New York (1997)
27. Lindvall, T.: *Lectures on the Coupling Method*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Wiley, New York (1992)
28. Maubon, J.: Riemannian symmetric spaces of the non-compact type: differential geometry. Technical report. Course at Summer School “Géométrie à courbure négative où nulle, groupes discrets et rigidités”, l’Institut Fourier (2004). <http://www-fourier.ujf-grenoble.fr/sites/ifmaquette.ujf-grenoble.fr/files/Maubon.pdf>
29. Molchanov, S.A.: Diffusion processes and Riemannian geometry. *Russ. Math. Surv.* **30**(1), 1 (1975)
30. Myers, S.B., Steenrod, N.: The group of isometries of a Riemannian manifold. *Ann. Math.* 400–416 (1939)
31. Nelson, E.: *Dynamical Theories of Brownian Motion*, 2nd edn. Mathematical Notes. Princeton University Press, Princeton (1967)
32. Petersen, P.: *Riemannian Geometry* (2006. Corr ed.). Graduate Texts in Mathematics. Springer, New York (2006)
33. Pitman, J.W.: On coupling of Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **35**(4), 315–322 (1976)
34. Protter, M.H., Weinberger, H.F.: *Maximum Principles in Differential Equations*. Springer, New York (1984). (Corrected reprint of the 1967 original)
35. Rogers, L.C.G.: Fastest coupling of random walks. *J. Lond. Math. Soc. (Second Series)* **60**(2), 630–640 (1999)
36. Sharan, P.: *Spacetime, Geometry and Gravitation*, Progress in Mathematical Physics, vol. 56. Springer, New York (2009)
37. Sverchkov, M.Y., Smirnov, S.N.: Maximal coupling for processes in  $D[0, \infty]$ . *Dokl. Akad. Nauk SSSR* **311**(5), 1059–1061 (1990)
38. Thorisson, H.: *Coupling, Stationarity, and Regeneration*. Springer, New York (2000)
39. Varadhan, S.R.S.: Diffusion processes in a small time interval. *Commun. Pure Appl. Math.* **20**(4), 659–685 (1967)
40. Varadhan, S.R.S.: Large deviations and applications. In: *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Montpelier (1984)