Coupling, local times, immersions

WILFRID S. KENDALL

Department of Statistics, University of Warwick, Coventry CV5 6FQ, UK.
E-mail: w.s.kendall@warwick.ac.uk.com; url: www.warwick.ac.uk/wsk

This paper answers a question of Émery [In *Séminaire de Probabilités XLII* (2009) 383–396 Springer] by constructing an explicit coupling of two copies of the Beneš *et al.* [In *Applied Stochastic Analysis* (1991) 121–156 Gordon & Breach] diffusion (BKR diffusion), neither of which starts at the origin, and whose natural filtrations agree. The paper commences by surveying probabilistic coupling, introducing the formal definition of an *immersed coupling* (the natural filtration of each component is immersed in a common underlying filtration; such couplings have been described as *co-adapted* or *Markovian* in older terminologies) and of an *equi-filtration* coupling (the natural filtration of each component is immersed in the filtration of the other; consequently the underlying filtration is simultaneously the natural filtration for each of the two coupled processes). This survey is followed by a detailed case-study of the simpler but potentially thematic problem of coupling Brownian motion together with its local time at 0. This problem possesses its own intrinsic interest as well as being closely related to the BKR coupling construction. Attention focusses on a simple immersed (co-adapted) coupling, namely the reflection/synchronized coupling. It is shown that this coupling is optimal amongst all immersed couplings of Brownian motion together with its local time at 0, in the sense of maximizing the coupling probability at all possible times, at least when not started at pairs of initial points lying in a certain singular set. However numerical evidence indicates that the coupling is *not* a maximal coupling, and is a simple but non-trivial instance for which this distinction occurs. It is shown how the reflection/synchronized coupling can be converted into a successful equi-filtration coupling, by modifying the coupling using a deterministic time-delay and then by concatenating an infinite sequence of such modified couplings. The construction of an explicit equi-filtration coupling of two copies of the BKR diffusion follows by a direct generalization, although the proof of success for the BKR coupling requires somewhat more analysis than in the local time case.

Keywords: bang-bang control; BKR diffusion; Brownian motion; co-adapted coupling; equi-filtration coupling; coupling; excursion theory; filtration; immersed coupling; Lévy transform; local time; Markovian coupling; maximal coupling; optimal immersed coupling; reflection coupling; reflection/synchronized coupling; stochastic control; synchronized coupling; Tanaka formula; Tanaka SDE; value function

1. Introduction

We begin with a brief survey of probabilistic coupling, which serves both to introduce some key concepts and to establish a context for the results proved in this paper. The concept of coupling has a long and distinguished history, dating back to Doeblin [11] (a biographical appreciation is given by Lindvall [27]). The method is now the subject of two scholarly expositions (Lindvall [28], Thorisson [40]), and has become a standard tool of the working probabilist (a somewhat more general concept appears in ergodic theory as the notion of a "joining"). Historically the thematic problem for coupling is that of constructing two copies of a given process on the same

sample space, starting at two different starting points but eventually coinciding. Such a coupling is said to be *successful*. In fact, many applications of coupling do not address the objective of eventually coinciding; nevertheless the thematic problem has been formative for the theory and remains significant in developing methods and intuition. Probabilistic coupling in general has found application throughout probability, for example in construction of gradient estimates, in distributional approximation (for instance, Stein–Chen approximation), in perfect simulation, and in monotonicity results for heat equations in insulated domains. The study of coupling in its own right is therefore a foundational topic for probability theory.

A landmark development in the study of coupling was the introduction of the notion of *maximal coupling*: a coupling which simultaneously maximises the chances of succeeding before time t for all possible t. Perhaps it will surprise the reader to learn that maximal couplings always exist: this was established by Griffeath [15] for time-homogeneous discrete Markov chains and by Goldstein [14] for more general discrete-time processes, based on a tail σ-algebra condition. (Note that even a maximal coupling need not necessarily have probability 1 of succeeding!) See also the very explicit construction given by Pitman [31] for time-homogeneous discrete Markov chains, Sverchkov and Smirnov's [38] note on coupling for continuous time using the J_1 topology, and Thorisson's [39] notion of *shift coupling*, which weakens the coupling requirement by allowing for general time-shifting of the coupled processes. (An informative treatment of some subtleties is given in the treatment of "faithful coupling" in Rosenthal [35].)

In general, the construction of maximal couplings is a demanding business: for substantial applications the task of construction is liable to require at least as much knowledge of the process in question as might be needed to solve the original problem to which the coupling method is to be applied. (Notwithstanding this general and justifiable pessimism, the simple *reflection coupling* of Brownian motion is a successful maximal coupling. Attention was originally drawn to this construction by the influential unpublished preprint of Lindvall [26].) More commonly, one works with less powerful couplings that are more easily constructed and analyzed, such as "co-adapted couplings". Co-adapted couplings (sometimes also called "Markovian couplings" in the context of coupling of Markov processes) require the two copies of the processes concerned to be adapted to the same filtration, and to have the same conditional laws based on conditioning on filtration σ-algebras. In the succinct language of filtrations (cf. Beghdadi-Sakrani and Emery [2], Émery [12,13]), the natural filtrations of the two processes must both be *immersed* in a common filtration (that is, the martingales of the natural filtrations must remain martingales in the larger common filtration). We therefore propose and adopt the new terminology of *immersed couplings* to replace the nomenclature of co-adapted or Markovian couplings:

Definition 1. *Consider two processes X and Y. An* immersed coupling *of X and Y is a construction of copies $\hat{X}$, $\hat{Y}$ of X, Y, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and adapted to the same filtration $\{\mathcal{F}_t : t \geq 0\}$, such that any martingale in the natural filtration of $\hat{X}$ remains a martingale in the common filtration $\{\mathcal{F}_t : t \geq 0\}$, and likewise for any martingale in the natural filtration of $\hat{Y}$.*

The extent to which immersed couplings are less powerful than maximal couplings was assessed in a preliminary way by Burdzy and Kendall [5], where they were studied in the guise

of Markovian couplings. As part of a study of *shy coupling* (the antithesis of the thematic coupling problem, in which one seeks to construct coupled copies which almost surely stay at least a fixed positive distance apart), Kendall [21], Lemma 6, records a characterization of immersed couplings of Brownian motion which has long been part of the general folklore of stochastic calculus: any immersed coupling of two d-dimensional Brownian motions A and B can be represented by the stochastic differential equation

$$\mathrm{d}A = J^\top \mathrm{d}B + K^\top \mathrm{d}C, \tag{1}$$

where C is a Brownian motion independent of B (perhaps to be defined after augmenting the filtration, if this is necessary to construct C), and J and K are two $(d \times d)$ matrix-valued predictable random processes satisfying $J^\top J + K^\top K = \mathbb{I}$ where $\mathbb{I}$ is the $(d \times d)$ identity matrix. We can view J as a predictable matrix-valued control for a somewhat degenerate stochastic control problem. (An informal discussion of links between stochastic control and coupling can be found in Kendall [20], Section 2.)

The terminology of immersed couplings is useful not only for its succinct definition, but also because it draws attention to a stricter constraint. Additionally, one could demand that either of the coupled copies could be constructed from the other, which corresponds to the requirement that the coupling possesses the *equi-filtration* property:

Definition 2. *Consider two processes X and Y. An* equi-filtration coupling *of X and Y is an immersed coupling $\hat{X}$, $\hat{Y}$ such that the natural filtration of $\hat{X}$ is equal to that of $\hat{Y}$.*

Of course it is the case that the equi-filtration coupling property follows from each natural filtration being immersed in the other. Consider one of the simplest nontrivial examples of coupling; Lindvall's Brownian reflection coupling is not only a successful maximal coupling, but also immersed and even equi-filtration. This is a very special case; for example Connor [10], Ph.D. thesis, considers reflection coupling of the Ornstein–Uhlenbeck process, if one copy of the Ornstein–Uhlenbeck process is started from 0 and the other copy is started from equilibrium. He notes that reflection coupling of the driving Brownian motions is clearly immersed but is not maximal even in the simple case. (Further exploration of the difference between maximality and immersion for couplings can be found in Kuwada and Sturm [25], Kuwada [24].)

The first objective of this paper is to investigate and explore properties of the construction of immersed and equi-filtration couplings in the simple case of coupling Brownian motion together with local time at 0. As a coupling problem this is only a little more complicated than the basic Brownian motion case, but it produces an example of existence of a successful immersed coupling (the reflection/synchronized coupling, Definition 3) which is optimal among all immersed couplings but (according to numerical evidence) is not maximal (Theorems 8, 9, 12 below). The reader may wish to compare other work on optimal immersed couplings for random walks on the line, on hypercubes and on hypercomplete graphs (Rogers [33], Connor and Jacka [9], Connor [8]).

A significant motivation for this study arises from the consideration that the reflection coupling has been a model for a wide variety of more sophisticated immersed couplings. For example, reflection coupling has been generalized to the case of elliptic diffusions with smooth coefficients

(Lindvall and Rogers [29], Chen and Li [7]), and also to the case of Riemannian Brownian motion (Kendall [18]), in which case there are connections with curvature properties. More recently, coupling techniques have been extended to cover some cases of hypoelliptic diffusions (Ben Arous *et al.* [3], Kendall and Price [23], Kendall [20,22]); essentially the issue here is to couple simultaneously not only Brownian motion but also one or more path functionals of the Brownian motion, namely time integrals, iterated time integrals, and Itô stochastic area integrals. Here it is necessary to augment the reflection coupling strategy with other coupling strategies, notably synchronous coupling and rotation coupling. In the stochastic differential framework (1), synchronous coupling corresponds to $K = 0$ and $J = \mathbb{I}$, while rotation coupling corresponds to $K = 0$ and J equal to a d-dimensional rotation. (It is interesting to compare this direction of research with the work of Émery [12], Theorem 1; this characterizes Brownian filtrations using the notion of "self-coupling" – jointly immersed Brownian filtrations for which a prescribed scalar functional is approximately coupled.)

While Brownian motion together with local time at 0 does not form a hypoelliptic diffusion in the strict sense, nevertheless the question of its coupling theory is clearly related to the hypoelliptic couplings mentioned above. The successful reflection/synchronized coupling is not only simple, but also (in view of the results proved here) evidently the right coupling for this situation. It is reasonable to hope that a careful and complete study of the reflection/synchronized coupling will be helpful in formulating and studying coupling methods for more general situations, as well as suggestive for coupling theory for hypoelliptic diffusions. The second objective of the paper is to demonstrate the first fruits of this aspiration and is fulfilled in Theorem 18 below, exhibiting a successful equi-filtration coupling for the BKR diffusion. The approach follows closely the methods developed for the reflection/synchronized coupling for Brownian motion together with local time at 0.

In summary, then, this paper conducts a case study of an almost surely successful coupling of a simple non-elliptic diffusion in the context of immersed and equi-filtration couplings; namely the reflection/synchronized coupling for Brownian motion together with local time at 0. The results of this case study are then applied to answer a question raised by Émery [13], by constructing an explicit equi-filtration coupling for BKR diffusions neither of which are begun at the origin.

Section 2 introduces the simple reflection/synchronized coupling for Brownian motion together with local time at 0, exploiting Tanaka's formula and the Lévy transform to re-cast the problem in terms of coupling Brownian motion together with a variant of its running supremum. The simplicity of this coupling allows for explicit calculation: in particular it is shown that the reflection/synchronized coupling is optimal amongst all immersed couplings, at least when their starting conditions are non-singular (here "optimal" means optimal in the sense of maximizing the probability of coupling by a given time t, for all possible times t, while "singular" means that in the re-cast form the two running suprema processes do not start from the same level). The moment-generating function for the coupling time is computed, and compared numerically with the moment-generating function for the maximal coupling time: numerical calculation then indicates that the reflection/synchronized coupling cannot be a maximal coupling.

The reflection/synchronized coupling is an immersed coupling but is not equi-filtration. Section 3 shows that if the couplings are perturbed by a simple deterministic time delay then it is possible to use a sequence of the resulting approximate couplings to construct a successful equi-filtration coupling of Brownian motion together with its local time at 0.

Section 4 introduces the BKR diffusion, sketches the immersed coupling described in Émery [13] (which bears a strong family resemblance to the reflection/synchronized coupling of Section 2, and which therefore is described here as a variant reflection/synchronized coupling), and notes that significant components of this variant reflection/synchronized coupling are actually immersed in the natural filtrations of both coupled diffusions. This is used to generate a successful equi-filtration coupling using the strategy of Section 3, hence answering Émery's question.

The paper is concluded by Section 5, which reviews the results of the paper and discusses some further research questions.

2. Coupling Brownian motion together with local time

The purpose of this section is to exhibit a successful immersed (but *not* equi-filtration) coupling for the two-dimensional diffusion made up of Brownian motion together with local time at zero. The simple construction (known already to Émery [13]) is based on Tanaka's formula for Brownian local time, and permits informative exact computations. In particular we are able to prove optimality of this coupling amongst all immersed couplings (Theorem 8), so long as the initial conditions are non-singular in a manner to be explained below, and thus to establish the optimal rate of immersed coupling (Theorem 9). We note in passing that this notion of optimality is distinct from the notion of ρ-optimality introduced by Chen [6].

2.1. Representation via the Tanaka formula

Recall the Tanaka formula or Lévy transform, expressing Brownian local time at 0 in terms of a stochastic integral:

$$\mathrm{d}|X| = \mathrm{sgn}(X)\,\mathrm{d}X + \mathrm{d}L^{(0)}. \tag{2}$$

Here X is a real Brownian motion and $L^{(0)}$ is the local time accumulated by X at 0. An immediate consequence of (2) is Lévy's famous transform, which represents $|X|$ and $L^{(0)}$ in terms of a new real Brownian motion B and S, a variant on the running supremum of B:

$$\begin{aligned} B &= L^{(0)} - |X|, \\ S &= L^{(0)}. \end{aligned} \tag{3}$$

It follows from (3) that $B = L_0^{(0)} - |X_0| - \int \mathrm{sgn}(X)\,\mathrm{d}X$ and $S_t = \max\{L_0^{(0)}, \sup\{B_s\colon s \le t\}\}$, so S does not start at B_0 if $|X_0| > 0$.

Evidently it suffices to exhibit successful coupling strategies for (B, S); off the line $X = 0$, this Lévy transform forms a $2\colon 1$ representation of the original pair $(X, L^{(0)})$; the two pre-images under the Lévy transform meet together when the Brownian motion X hits 0.

2.2. The reflection/synchronized coupling for immersed coupling of Brownian motion together with local time

The above considerations show that it suffices to exhibit a successful immersed coupling between (a) the pair (B, S) above and (b) a copy $(\widetilde{B}, \widetilde{S})$ started with different initial conditions. Were the corresponding X and $\widetilde{X}$ not to agree at coupling, one could simply continue with synchronized coupling until $|X| = |\widetilde{X}|$ hits 0. However, the reflection/synchronized coupling given below actually terminates with $B = S$ and $\widetilde{B} = \widetilde{S}$, so at the end of this coupling we already have $|X| = |\widetilde{X}| = 0$. Without loss of generality, suppose that $B_0 = L_0^{(0)} - |X_0| \geq \widetilde{B}_0 = \widetilde{L}_0^{(0)} - |\widetilde{X}_0|$.

Definition 3 (Reflection/synchronized coupling). *The* reflection/synchronized coupling *algorithm consists of two stages*:

1. Reflection coupling $(\mathrm{d}B = -\mathrm{d}\widetilde{B})$ *till the time* $T_1 = \inf\{t: B_t = \widetilde{B}_t\}$ *(the first time that B and $\widetilde{B}$ meet)*; *then (if (B, S) is not already coupled with $(\widetilde{B}, \widetilde{S})$).*
2. Synchronized coupling $(\mathrm{d}B = +\mathrm{d}\widetilde{B})$, *run from time T_1 until the time $T_2 = \inf\{t > T_1: B_t \equiv \widetilde{B}_t = S_{T_1} \vee \widetilde{S}_0\}$ that $B \equiv \widetilde{B}$ first hits the higher level $S_{T_1} \vee \widetilde{S}_0$ after time T_1.*

Note that at the end of stage 2 we have $B = S$ and $\widetilde{B} = \widetilde{S}$, so $|X| = |\widetilde{X}| = 0$.

It is possible for the coupling of (B, S) and $(\widetilde{B}, \widetilde{S})$ to be abbreviated to a one-stage (reflection) coupling in case $S_0 = \widetilde{S}_0$, for if it happens that B (and therefore $\widetilde{B}$) both stay below $S_0 = \widetilde{S}_0$ up to time T_1 then coupling will be successfully achieved at time $T_{\mathrm{couple}} = T_1 < T_2$. However we will see below that this case can be viewed as singular, as a consequence of Lemma 6. Moreover if X_0 and $\widetilde{X}_0$ are of opposite sign then they will not couple at this stage: it still will be necessary to proceed to completion of the synchronization stage so that $B_{T_2} = S_{T_2} = \widetilde{B}_{T_2} = \widetilde{S}_{T_2}$ and therefore $X_{T_2} = \widetilde{X}_{T_2} = 0$.

If not completed at the end of the reflection stage, then the coupling will succeed at the time $T_{\mathrm{couple}} = T_2$; at that moment in time it is the case that simultaneously $B = \widetilde{B}$ (since they are coupled by synchronization after meeting at time T_1) and $S = \widetilde{S}(= B = \widetilde{B})$. Note that the $[T_1, T_2]$ stage depends on the behaviour of B over the initial time interval $[0, T_1]$. Indeed, note that by construction (and particularly by choice of initial conditions $B_0 = L_0^{(0)} - |X_0| \geq \widetilde{B}_0 = \widetilde{L}_0^{(0)} - |\widetilde{X}_0|$) it is the case that $\widetilde{B}$ will stay below or equal to B until time T_2, and hence $S_{T_1} \vee \widetilde{S}_0 = S_{T_1} \vee \widetilde{S}_{T_1}$. The construction is illustrated in Figure 1.

The coupling is almost surely successful, since both the first and second stages correspond to times taken for real Brownian motion to hit specified levels. Indeed, the coupling of (B, S) and a copy $(\widetilde{B}, \widetilde{S})$ is equi-filtration, not just immersed, because the stopping times T_1 and T_2 can be rewritten as hitting times for B in its natural filtration (successively, from B_0 to $\frac{1}{2}(B_0 + \widetilde{B}_0)$), then from $\frac{1}{2}(B_0 + \widetilde{B}_0)$ to $S_{T_1} \vee \widetilde{S}_0$), and similarly also as hitting times for $\widetilde{B}$ in its own natural filtration. (In particular, T_1 can be rewritten as the hitting time of $\widetilde{B}$ moving from $\widetilde{B}_0$ to $\frac{1}{2}(B_0 + \widetilde{B}_0)$.) The corresponding immersed coupling of $(X, L^{(0)})$ with $(\widetilde{X}, \widetilde{L}^{(0)})$ cannot be an equi-filtration coupling, because the natural filtration of X has to be augmented in order to supply appropriate randomness for the signs of the excursions of $\widetilde{X}$ from zero. (See Émery [13], Lemma 5, for a similar augmentation in the more complicated case of BKR diffusions.)
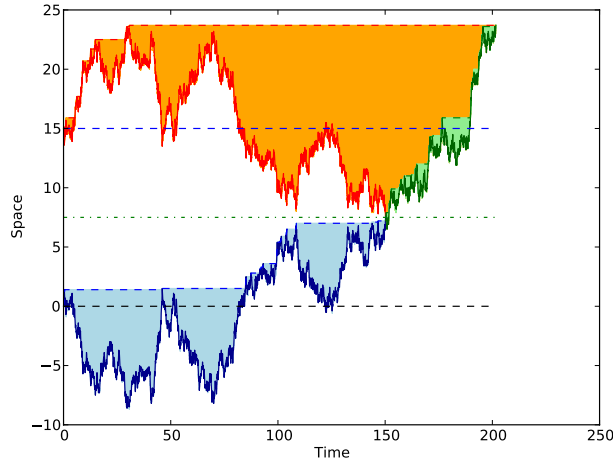
Figure 1. Illustration of a successful reflection/synchronized coupling of Brownian motion B together with S.

There is a natural reformulation of reflection/synchronized coupling in terms of stochastic calculus: set $\mathrm{d}\widetilde{B} = J\,\mathrm{d}B$ up to the coupling time T_2, where the predictable control J is given very simply by

$$J_t = \begin{cases} -1 & \text{for } t < T_1 \text{ (reflection stage),} \\ +1 & \text{for } T_1 \leq t \leq T_2 \text{ (synchronized stage).} \end{cases} \tag{4}$$

The failure of mutual immersion for the coupling of $(X, L^{(0)})$ with $(\widetilde{X}, \widetilde{L}^{(0)})$ is immediately apparent from the relevant stochastic differential equation

$$\mathrm{d}\widetilde{X} = \mathrm{sgn}(\widetilde{X})\,J\,\mathrm{sgn}(X)\,\mathrm{d}X, \tag{5}$$

which is an instance of Tanaka's classic example of a Brownian motion $\widetilde{X}$, defined as a weak but not strong solution of a stochastic differential equation driven by a second Brownian motion $\int J\,\mathrm{sgn}(X)\,\mathrm{d}X$.

2.3. Optimality amongst immersed couplings

The reflection/synchronized coupling strategy is faster than all other immersed couplings, in the sense that it minimizes

$$\mathbb{P}[T_{\mathrm{couple}} > t]$$

simultaneously for all $t > 0$, except perhaps for the singular case of $S_0 = \widetilde{S}_0$ (this singular case is discussed around the statement of Lemma 6 below). Equivalently the distribution of the coupling time T_{couple} for any immersed coupling exhibits stochastic domination over the distribution of T_{couple} for the reflection/synchronized coupling (except perhaps in singular cases).

Before stating and proving a theorem which asserts this optimality, we first establish some preparatory lemmas. The first one concerns the coupling of two Brownian motions on $[0, \infty)$ that are stopped when the first one of them hits 0.

Lemma 4. *Suppose the planar process (U, V) is composed of two Brownian motions which are related by an immersed coupling, and suppose that (U, V) is started at a point (U_0, V_0) in the interior of the quadrant $\{(u, v): u \geq 0, v \geq 0\}$. Let T be the first time that (U, V) hits the boundary of the quadrant. Suppose it is desired to construct the coupling so that $\mathbb{P}[(U_T, V_T) = (0, 0)] = 1$. This is possible if and only if $U_0 = V_0$ and the coupling is the synchronized coupling.*

Proof. Using the formalism of Itô [17] (see also Ikeda and Watanabe [16], Chapter III.1, and the development in Kendall [19]) and the representation of immersed Brownian couplings given in (1), the general law of (U, V) under an immersed coupling produces $dU^2 = dV^2 = dt$, Drift $dU =$ Drift $dV = 0$, and $dU\, dV = J\, dt$ for an arbitrary adapted integrand $J \in [-1, 1]$ which can be viewed as the control for the stochastic control problem of maximizing the objective function $\mathbb{P}[(U_T, V_T) = (0, 0)]$.

Without loss of generality, we may suppose that $U_0 \geq V_0 > 0$. Note that under reflection coupling ($J = -1$) the probability of (U, V) hitting the diagonal $\{(u, v): u = v\}$ before time T is given by

$$\Phi(U_0, V_0) = \frac{V_0}{(1/2)(U_0 + V_0)}.$$

We extend the definition of Φ to the case $V_0 \geq U_0 > 0$ by setting $\Phi(U_0, V_0) = \Phi(V_0, U_0)$, so that

$$\Phi(U, V) = \min\left\{\frac{U}{(1/2)(U + V)}, \frac{V}{(1/2)(U + V)}\right\}.$$

An application of Itô calculus shows that if $U > V > 0$ then, under a general control $J \in [-1, 1]$,

$$\text{Drift}\, d\Phi(U, V) = -\frac{2}{(U + V)^3}(U - V)(1 + J)\, dt,$$

and this is non-positive, and vanishes only when $J = -1$. A similar result holds for $V > U > 0$. On the other hand, if $U_0 = V_0 > 0$ and $J = +1$ then (U, V) stays on the diagonal, so that $\Phi(U, V)$ then remains constant. An argument using the Itô–Tanaka formula for semimartingales thus shows that $\Phi(U, V)$ is a supermartingale for all immersed couplings of U and V, and becomes a martingale only under the strategy "use reflection coupling till (U, V) hits the diagonal or the boundary, then use synchronized coupling till (U, V) hits the boundary". It follows that $\Phi(U_0, V_0)$ is the maximum of $\mathbb{P}[(U_T, V_T) = (0, 0)]$ over all immersed couplings, and is attained only by using this strategy. The lemma follows. $\qquad\square$

As a consequence of the lemma, we can prove the optimality of the reflection/synchronized coupling in the special case when $B_0 = \widetilde{B}_0$. This allows us to restrict attention to immersed couplings which preserve the ordering of B and $\widetilde{B}$.

Lemma 5. *Consider the reflection/synchronized coupling (Example 3) for the special case $B_0 = \widetilde{B}_0$. This is the only optimal coupling amongst all immersed couplings started with $B_0 = \widetilde{B}_0$, and so (since the reflection stage succeeds immediately) the uniquely optimal way to proceed is to cease immediately if $S_0 = \widetilde{S}_0$, and otherwise to conduct a synchronized coupling of B and $\widetilde{B}$ until $B \equiv \widetilde{B}$ hits $S_0 \vee \widetilde{S}_0$.*

Proof. If $S_0 = \widetilde{S}_0$, then coupling succeeds immediately and there is nothing to prove. Suppose without loss of generality that $S_0 > \widetilde{S}_0$. Coupling cannot succeed earlier than the first time $\widetilde{T}$ at which $\widetilde{B}$ hits S_0, and if we employ synchronized coupling then coupling will succeed at this hitting time.

This shows that synchronized coupling is optimal, but we require strict optimality. Consider a second coupling which does not employ synchronized coupling throughout. It then follows that there must be a moment, *before* time $\widetilde{T}$, at which either $B > \widetilde{B}$ or $\widetilde{B} > B$. We can apply Lemma 4 to $U = S_0 - B$ and $V = S_0 - \widetilde{B}$; it follows that if synchronized coupling is not employed right up to time $\widetilde{T}$, then there is a positive probability that one of two possible cases has occurred: either B has already hit S_0 by time $\widetilde{T}$, or B has not yet hit S_0 by time $\widetilde{T}$. In the first case, the properties of Brownian motion B show that almost surely $S_{\widetilde{T}} > S_0$, and so successful coupling must occur after $\widetilde{B}$ travels from S_0 to $S_{\widetilde{T}}$. In the second case, successful coupling must wait at least until $\widetilde{B}$ and B meet after time $\widetilde{T}$.

It follows that, for any coupling other than synchronized coupling, (a) the coupling time can be no less than $\widetilde{T}$, (b) there is a positive chance of it being strictly greater than $\widetilde{T}$. This establishes the required stochastic domination (since $\widetilde{T}$ is the hitting time of a real Brownian motion started at $B_0 = \widetilde{B}_0$ and rising to $S_0 > S_0 \vee \widetilde{S}_0$) and so the lemma follows.   □

In passing, we are now able to explain the reason why it is appropriate to describe as singular the case when $B_0, \widetilde{B}_0 < S_0 = \widetilde{S}_0$. In this case it is possible for full coupling of (B, S) with $(\widetilde{B}, \widetilde{S})$ to succeed as soon as B first meets $\widetilde{B}$, so long as B and $\widetilde{B}$ do not hit $S_0 = \widetilde{S}_0$ (as noted in Section 2.2, this need not imply success of the coupling of X with $\widetilde{X}$ at that time). The next lemma shows that if $S_0 \neq \widetilde{S}_0$ then this early success cannot occur.

Lemma 6. *Suppose $S_0 \neq \widetilde{S}_0$. Then an optimal immersed coupling of (B, S) and $(\widetilde{B}, \widetilde{S})$ succeeds exactly at the first time when B, S, $\widetilde{B}$, and $\widetilde{S}$ simultaneously coincide.*

Proof. Consider first the case $B_0 = \widetilde{B}_0$. As shown by Lemma 5, it is then the case that the only optimal immersed coupling is provided by synchronized coupling until $B = \widetilde{B}$ first hits $S_0 \vee \widetilde{S}_0$, and the characterization of coupling by simultaneous coincidence is immediate.

Consider the case $\widetilde{B}_0 < B_0$ (the case of $\widetilde{B}_0 < B_0$ is entirely similar). It is a consequence of Lemma 5 that optimality of the immersed coupling implies that the relationship $\widetilde{B} \leq B$ must persist till full coupling is successful.

So further suppose that $\widetilde{S}_0 < S_0$. In that sub-case it follows from $\widetilde{B} \leq B$ that the relationship $\widetilde{S} \leq S$ must persist till full coupling is successful. Coupling cannot succeed till $\widetilde{S}$ hits S, and when that happens we must have $\widetilde{B} = \widetilde{S}$. But in this sub-case we also have $\widetilde{B} \leq \widetilde{S} \leq S$ and $\widetilde{B} \leq B \leq S$. Consequently full coupling must succeed when $\widetilde{S}$ first hits S, and at that time B, S, $\widetilde{B}$, and $\widetilde{S}$ simultaneously coincide.

Suppose on the other hand that $S_0 < \widetilde{S}_0$. In that sub-case again, full coupling cannot succeed before S hits $\widetilde{S}$, at which time it is necessary that B also hits $\widetilde{S}$. If it is further the case that $\widetilde{B} = B$ at that time, then full coupling succeeds in the manner prescribed by the lemma. If on the other hand $\widetilde{B} < B$ at that time, then (by the properties of Brownian motion) there are instants immediately after this time at which $\widetilde{B} < \widetilde{S} < B < S$, and we can proceed as above. ☐

We shall now show that the distribution of the coupling time T_{couple} under any immersed coupling can be dominated in the limit (as $N \to \infty$) by the distribution of T_{couple} under an immersed coupling whose predictable control J is restricted to values ± 1 (thus, a "bang-bang" control), and moreover such that J is constant on stochastic intervals $[\tau_k^{(N)}, \tau_{k+1}^{(N)})$ defined as follows. For any positive even integer $N > 0$, consider the one-dimensional lattice $\mathcal{L}^{(N)}$ (depending implicitly on B_0, and $\widetilde{B}_0$)

$$\mathcal{L}^{(N)} = B_0 + \frac{\widetilde{B}_0 - B_0}{N} \mathbb{Z} = \left\{ B_0 + \frac{k}{N}(\widetilde{B}_0 - B_0): \ k = 0, \pm 1, \pm 2, \dots \right\}.$$

We define a *mesh*, a sequence of stopping times $0 = \tau_0^{(N)} < \tau_1^{(N)} < \tau_2^{(N)} < \cdots$, as a sequence of "crossing times" for this lattice:

$$\tau_{k+1}^{(N)} = \inf\left\{ t > \tau_k^{(N)}: \ B_t \in \mathcal{L}^{(N)} \setminus \{B_{\tau_k^{(N)}}\} \right\}.$$

Sampling using this mesh of stopping times has the effect of discretizing the Brownian motion B into a random walk with steps $\pm \frac{1}{N}(\widetilde{B}_0 - B_0)$.

Note, for an immersed coupling restricted to a control J which is locally constant with $J = \pm 1$ on each stochastic interval $[\tau_k^{(N)}, \tau_{k+1}^{(N)})$ of the mesh:

1. both B and $\widetilde{B}$, when sampled at times $0 = \tau_0^{(N)} < \tau_1^{(N)} < \tau_2^{(N)} < \cdots$, belong to the lattice $\mathcal{L}^{(N)}$, since $\widetilde{B}$ is obtained from B using a predictable control J formed from synchronizations and reflections and which alters only when B belongs to the lattice;

2. because N is even, T_{couple} belongs to the set $\{\tau_0^{(N)}, \tau_1^{(N)}, \tau_2^{(N)}, \dots\}$;

3. finally, our candidate for optimality, the reflection/synchronized coupling (Example 3), can itself be viewed as one of these couplings, since the control changes from $+1$ to -1 exactly at one of the stopping times in the mesh. (This is the reason why it is convenient to work with meshes of stopping times, rather than decompositions of the time axis into disjoint dyadic intervals.)

We can now summarize and prove a result stating that an optimal immersed coupling can be approximated in distribution by appropriately chosen "bang-bang" controls of the above form. The proof is related to the method of proof of Émery [12], Proposition 2; however here we need the control J to have the "bang-bang" property rather than simply to be locally constant, and to be composed of stopping times drawn from a mesh of stopping times as specified above.

Lemma 7. *For any fixed $t > 0$, any optimal immersed coupling of (B, S) and $(\widetilde{B}, \widetilde{S})$ can be approximated weakly over $[0, t]$ (when viewed as a probability distribution on the metric space of*

4-dimensional continuous trajectories, equipped with the sup-norm) by "bang-bang" immersed couplings for which the control J takes values ±1 only, and only changes at hitting times belonging to some mesh.

Note that the lemma does *not* assert that the "bang-bang" couplings are successful!

Proof of Lemma 7. Consider a general immersed coupling determined by $\widetilde{B} = \widetilde{B}_0 + \int J \, dB$ and subject to the constraint that the coupling is synchronized once B and $\widetilde{B}$ have met. (By Lemma 5, all optimal immersed couplings must be of this form.) Since $|J| \leq 1$, for each $t > 0$ we have $\mathbb{E}[\int_0^t J^2 \, ds] < \infty$, and moreover for each $\varepsilon > 0$ we may find continuous predictable f with $\mathbb{E}[\int_0^t |f - J|^2 \, ds] < \varepsilon^2/4$ (for example, $f_t = \frac{2}{\delta^2} \int_{t-\delta}^t (t - s) J_s \, ds$ for sufficiently small δ). It then follows that for sufficiently large N we may approximate f in L^2 by *piece-wise constant* $J^{[c]}$ such that $J^{[c]} \in [-1, 1]$ is predictably constant on each dyadic interval $[\tau_k^{(N)}, \tau_{k+1}^{(N)})$ of the mesh, and $\mathbb{E}[\int_0^t |f - J^{[c]}|^2 \, ds] < \varepsilon^2/4$, hence

$$\mathbb{E}\left[\int_0^t \left| J - J^{[c]} \right|^2 ds \right] < \varepsilon^2.$$

Doob's submartingale inequality then implies that we can control

$$\sup_{s \leq t}\left\{ \left| \int_0^s J \, dB - \int_0^s J^{[c]} \, dB \right| \right\},$$

so that $\widetilde{B}_0 + \int J^{[c]} \, dB$ is a good path-wise approximation to $\widetilde{B} = \widetilde{B}_0 + \int J \, dB$.

While $J^{[c]}$ is piecewise-constant on stochastic intervals related to the mesh, it does not take values in $\{\pm 1\}$. We need an approximation based on a "bang-bang" control $J^{[bb]}$, which is constrained by $J^{[bb]} \in \{\pm 1\}$ as well as by the requirement that $J^{[bb]}$ is predictably constant on stochastic intervals $[\tau_k^{(M)}, \tau_{k+1}^{(M)})$ which now must be formed on a new mesh, defined for some still larger even integer $M = 2^r N$, for an integer $r > 0$. Given $M > N$, we define $J^{[bb];(M)}$ to "track" $J^{[c]}$ in the following co-adapted way:

$$J^{[bb];(M)}_{\tau_k^{(M)}} = \begin{cases} +1 & \text{if } \int_0^{\tau_k^{(M)}} J^{[bb];(M)} \, du \leq \int_0^{\tau_k^{(M)}} J^{[c]} \, du, \\ -1 & \text{if } \int_0^{\tau_k^{(M)}} J^{[bb];(M)} \, du > \int_0^{\tau_k^{(M)}} J^{[c]} \, du. \end{cases} \tag{6}$$

Since $|J| \leq 1$, it follows that we have the following bound for $s \in [0, t]$:

$$\left| \int_0^s J^{[c]} \, du - \int_0^s J^{[bb];(M)} \, du \right| \leq 2 \sup\{ (\tau_{k+1}^{(M)} \wedge t) - (\tau_k^{(M)} \wedge t) : k = 1, 2, \ldots \}, \tag{7}$$

converging almost surely to zero as $2^r = M/N \to \infty$.

Consider the sequence of two-dimensional processes $\{(B, \widetilde{B}_0 + \int J^{[bb];(M)} \, dB) : M = 2^r N\}$, defined on the time-range $[0, t]$. The one-dimensional coordinate processes being Brownian motions, it follows that this sequence is tight. Any convergent subsequence converges to a limit for

which the one-dimensional coordinate processes are Brownian motions; moreover, using (7), we may deduce that in the limit the product of the pair of one-dimensional coordinate processes is equal to the sum of a martingale and the integral $\int_0^s J^{[c]} \, du$. Hence by semimartingale Itô calculus the limit has the law of $(B, \widetilde{B}_0 + \int J^{[c]} \, dB)$, no matter what convergent subsequence is chosen, and therefore by the theory of weak convergence we may deduce that the sequence of random paths $(B, \widetilde{B}_0 + \int J^{[bb];(M)} \, dB)$ converges weakly to this limit.

It follows that we can choose a sequence of "bang-bang" controls $J^{(n)}$, constant on appropriate meshes $\{[\tau_k^{(M_n)}, \tau_{k+1}^{(M_n)}): k = 1, 2, \ldots\}$ (with $M_n \to \infty$), such that $(B, \widetilde{B}_0 + \int J^{(n)} \, dB)$ converges weakly (using supremum norm over the time interval $[0, t]$) to the immersed coupling $(B, \widetilde{B})$ which was originally under consideration. $\qquad \square$

We can now argue for optimality of the reflection/synchronized coupling in the general non-singular case ($S_0 \neq \widetilde{S}_0$). We need only consider the case when $B_0 \neq \widetilde{B}_0$, since Lemma 5 covers the case of $B_0 = \widetilde{B}_0$; indeed it suffices to consider only those immersed couplings which are constrained to be synchronized couplings once B and $\widetilde{B}$ have met. Employing the terminology of Section 2.2, we set $T_1 = \inf\{t : B_t = \widetilde{B}_t\}$. Thus, we need consider only those immersed couplings for which $J_t = 1$ once $t > T_1$.

Theorem 8. *Suppose that $B_0 \neq \widetilde{B}_0$ and $S_0 \neq \widetilde{S}_0$. The reflection/synchronized coupling* (*Example* 3) *is optimal amongst all immersed couplings of Brownian motion together with local time.*

Proof. As noted above, by Lemma 5 we may restrict attention to immersed couplings for which (without loss of generality) $B \geq \widetilde{B}$, and such that $B \equiv \widetilde{B}$ after $T_1 = \inf\{s : B_s = \widetilde{B}_s\}$. Moreover, by the argument of Lemma 6, at the coupling time T_{couple} we must have $B_{T_{\text{couple}}} = S_{T_{\text{couple}}} = \widetilde{B}_{T_{\text{couple}}} = \widetilde{S}_{T_{\text{couple}}}$.

The first step is to use the weak approximations $(B, \widetilde{B}_0 + \int J^{(n)} \, dB)$ of $(B, \widetilde{B})$ (as given in Lemma 7) to build *successful* immersed couplings of (B, S) and $(\widetilde{B}, \widetilde{S})$ with coupling times which are in the limit stochastically dominated by the coupling time derived from $(B, \widetilde{B})$. For convenience, we employ the Skorokhod representation of weak convergence; augmenting the probability space if necessary, we construct a copy $(B^{(n)}, \widetilde{B}^{(n)} = \widetilde{B}_0 + \int J^{*,(n)} \, dB^{(n)})$ of $(B, \widetilde{B}_0 + \int J^{(n)} \, dB)$ on the same probability space as $(B, \widetilde{B})$ such that almost surely $B^{(n)} \to B$ and $\widetilde{B}^{(n)} \to \widetilde{B}$ uniformly on the time interval $[0, t]$. (We note in passing that this construction need not respect the underlying filtration. The stochastic integrand $J^{*,(n)}$ and the stochastic integral $\int J^{*,(n)} \, dB^{(n)}$ are defined with respect to the natural filtration of $B^{(n)}$, which need not immerse in the original filtration!)

Although the target coupling of (B, S) and $(\widetilde{B}, \widetilde{S})$ couples at T_{couple}, we should not suppose that $(B^{(n)}, S^{(n)})$ and $(\widetilde{B}^{(n)}, \widetilde{S}^{(n)})$ couple at this time (using $S^{(n)}$ and $\widetilde{S}^{(n)}$ to denote the corresponding supremum processes). However, we can modify $(B^{(n)}, \widetilde{B}^{(n)})$ to produce a coupling which does not succeed much later than the original coupling.

Indeed, we have restricted attention to immersed couplings such that at the coupling time T_{couple} we have $B_{T_{\text{couple}}} = S_{T_{\text{couple}}} = \widetilde{B}_{T_{\text{couple}}} = \widetilde{S}_{T_{\text{couple}}}$. Accordingly, we may choose a sequence $\varepsilon_n \to 0$ such that

$$\mathbb{P}\big[B_{T_{\text{couple}}}^{(n)}, S_{T_{\text{couple}}}^{(n)}, \widetilde{B}_{T_{\text{couple}}}^{(n)}, \widetilde{S}_{T_{\text{couple}}}^{(n)} \text{ all lie within } \pm \varepsilon_n \text{ of each other}\big] \geq 1 - \varepsilon_n.$$

Accordingly, if we set

$$T_3^{(n)} = \inf\{s: B_s^{(n)}, S_s^{(n)}, \widetilde{B}_s^{(n)}, \widetilde{S}_s^{(n)} \text{ all lie within } \pm \varepsilon_n \text{ of each other}\},$$

then

$$\mathbb{P}[T_3^{(n)} > t] \leq \mathbb{P}[T_{\text{couple}} > t] + \varepsilon_n.$$

But at time $T_3^{(n)}$ we can modify the construction of $(B^{(n)}, S^{(n)})$ and $(\widetilde{B}^{(n)}, \widetilde{S}^{(n)})$ to use the reflection/synchronized coupling (Example 3), obtaining successful coupling at time $T_{\text{couple}}^{(n)} \geq T_3^{(n)}$. As $\varepsilon \to 0$ so we can deduce by scaling that the extra time $T_{\text{couple}}^{(n)} - T_3^{(n)}$ required for success of this final coupling must tend to zero in probability.

It follows from these arguments that the infimum of the probability of failing to couple before time t, for any fixed $t > 0$,

$$\mathbb{P}[T_{\text{couple}} > t],$$

can be approached by considering $\mathbb{P}[T_{\text{couple}} > t + \varepsilon_n]$ for suitable $\varepsilon_n \to 0$ and immersed couplings based on "bang-bang" controls $J^{[bb]}$ constrained by change only at stopping times taken from meshes $0 = \tau_0^{(N)} < \tau_1^{(N)} < \tau_2^{(N)} < \cdots$, and which become synchronous after B and $\widetilde{B}$ first meet.

Consider such an immersed coupling with control $J^{[bb]}$. For a fixed $t > 0$, we consider the following value function, defined for $0 \leq u < t$:

$$
\begin{aligned}
V(u; b, \widetilde{b}, s, \widetilde{s}) \\
= \mathbb{P}[T_{\text{couple}} > t - u | B_u = b, \widetilde{B}_u = \widetilde{b}, S_u = s, \widetilde{S}_u = \widetilde{s}; \text{reflection/synchronized coupling}].
\end{aligned}
\tag{8}
$$

We are particularly interested in the discrete-time process obtained by sampling at stopping times taken from the specified mesh, but stopping at the terminal time t:

$$\{Z_n = V(\tau_n^{(N)} \wedge t; B_{\tau_n^{(N)} \wedge t}, \widetilde{B}_{\tau_n^{(N)} \wedge t}, S_{\tau_n^{(N)} \wedge t}, \widetilde{S}_{\tau_n^{(N)} \wedge t}): n = 0, 1, 2, \ldots\}.$$

It follows by definition that $V(u; B_u, \widetilde{B}_u, S_u, \widetilde{S}_u)$ is a bounded martingale under the reflection/synchronized coupling. Under this coupling Z is a discrete-time martingale since it is obtained from the bounded process $\{V_{u \wedge t}: u \geq 0\}$ by sampling at stopping times. We shall show that Z is a supermartingale under the coupling specified by $J^{[bb]}$, and moreover that the martingale property cannot hold if $J^{[bb]} = +1$ over the initial time interval $[0, \tau_1^{(N)})$. Arguing inductively, this suffices to establish the theorem.

The crux of the matter is to consider the behaviour of the value function at time zero if the initial segment of coupling is synchronized. To this end, we make a special construction of the reflection/synchronized coupling referred to by the value function: we suppose two independent Brownian motions are employed (both begun at 0), namely $B^{(r)}$ to drive the reflection stage of the coupling, and $B^{(s)}$ to drive the synchronized stage. We set

- $\tau_1^{(N;s)} = \inf\{t > 0: |B_t^{(s)}| = \frac{1}{N}(B_0 - \widetilde{B}_0)\}$,

- $T_1^{(r)}$ to be the time when $B^{(r)}$ first hits $-\frac{1}{2}(B_0 - \widetilde{B}_0)$, corresponding to the end of the reflection stage,
- and $M^{(r)} = \sup\{B_s^{(r)}: s \leq T_1^{(r)}\} + B_0$ to be the maximum level achieved during the reflection stage.

Then the reflection/synchronized coupling time corresponds in law to $T_3^{(*)} = T_1^{(r)} + \inf\{s: B_s^{(s)} + \frac{1}{2}(B_0 + \widetilde{B}_0) = \max\{S_0 \vee \widetilde{S}_0, M^{(r)}\}\}$, and so $Z_0 = \mathbb{P}[T_3^{(*)} > t]$.

Now consider the effect of commencing with a session of synchronized coupling. We consider two possible cases. Suppose in the first case that

$$M^{(r)} - \frac{1}{N}(B_0 - \widetilde{B}_0) \leq S_0 \vee \widetilde{S}_0.$$

Then we can represent the initial session of synchronized coupling by using $B^{(s)}|_{[0, \tau_1^{(N;s)})}$, and then replacing $B_t^{(s)}$ by $B_{t+\tau_1^{(N;s)}}^{(s)} - B_{\tau_1^{(N;s)}}^{(s)}$. Evidently the distribution of $T_3^{(*)}$ is unaffected by this change. If furthermore

$$M^{(r)} + \frac{1}{N}(B_0 - \widetilde{B}_0) \leq S_0 \vee \widetilde{S}_0$$

then the reflection stage will start at time $\tau_1^{(N;s)}$ at level $B_{\tau_1^{(N;s)}}^{(s)}$, moreover by the end of the reflection stage the supremum of the coupled processes will not exceed $S_0 \vee \widetilde{S}_0$, and the subsequent synchronization stage will have to move from $B_{\tau_1^{(N;s)}}^{(s)} + \frac{1}{2}(B_0 + \widetilde{B}_0)$ to $S_0 \vee \widetilde{S}_0$. It follows that $T_3^{(*)}$ is still the coupling time. If on the other hand

$$M^{(r)} + \frac{1}{N}(B_0 - \widetilde{B}_0) > S_0 \vee \widetilde{S}_0,$$

then there is a possibility that the supremum of the coupled processes *will* exceed $S_0 \vee \widetilde{S}_0$. However, the subsequent synchronization stage will still have to move from $B_{\tau_1^{(N;s)}}^{(s)} + \frac{1}{2}(B_0 + \widetilde{B}_0)$ to $S_0 \vee \widetilde{S}_0$, but may have to move even further. Thus, the coupling time still cannot occur earlier than $T_3^{(*)}$.

Suppose in the second case that

$$M^{(r)} - \frac{1}{N}(B_0 - \widetilde{B}_0) > S_0 \vee \widetilde{S}_0.$$

Then we can represent the initial session of synchronized coupling by using an independent copy $\hat{B}|_{[0, \hat{\tau}_1^{(N;s)})}$ of $B|_{[0, \tau_1^{(N;s)})}$, and restarting the construction at the new starting points $B_0 \pm \frac{1}{N}(B_0 - \widetilde{B}_0)$, $\widetilde{B}_0 \pm \frac{1}{N}(B_0 - \widetilde{B}_0)$ (same sign for each initial increment). Regardless of the sign of the initial increment, coupling occurs at $\hat{\tau}_1^{(N;s)} + T_3^{(*)}$, so is delayed relative to $T_3^{(*)}$.

It follows from these arguments that an initial session of synchronized coupling followed by reflection/synchronized coupling cannot increase the probability of successful coupling by time t compared with that of reflection/synchronized coupling, and has a positive chance of reducing it; consequently Z is a supermartingale and the martingale property for Z cannot hold if $J^{[bb]} = +1$ over the initial time interval $[0, \tau_1^{(N)})$.

Since Z is a martingale for the reflection/synchronized coupling, this suffices to establish the theorem. $\square$

It is likely that the reflection/synchronized coupling is the *unique* optimal immersed coupling, but we do not pursue this technicality here.

In the singular case the value function (8) takes on a more complicated form, and the above approach in this case will no longer settle whether or not the reflection/synchronized coupling is optimal amongst immersed couplings.

2.4. Rate of optimal immersed coupling

We now elicit the rate at which the reflection/synchronized coupling occurs. This is accomplished by calculating the moment generating function of the optimal immersed coupling time in the non-singular case of $S_0 \neq \widetilde{S}_0$, supposing (without loss of generality) that $B_0 > \widetilde{B}_0$.

Theorem 9. *Let T_{couple} be the coupling time for Brownian motion together with local time, equivalently for Brownian motion B together with its supremum S, using the reflection/synchronized coupling described above. Let $\widetilde{B}$ and $\widetilde{S}$ be the corresponding coupled quantities. Under the non-singular conditions $S_0 = B_0 > \widetilde{B}_0 = \widetilde{S}_0$, the coupling time has the following moment generating function:*

$$\mathbb{E}\big[\exp(-\alpha T_{\text{couple}})\big] = 1 + \sinh\left(\sqrt{\frac{\alpha}{2}}(B_0 - \widetilde{B}_0)\right) \log \tanh\left(\sqrt{\frac{\alpha}{2}} \frac{B_0 - \widetilde{B}_0}{2}\right).$$

Proof. Recall the notation of Section 2.2, and bear in mind the stipulation that $S_0 \neq \widetilde{S}_0$. It is required to calculate the moment generating function

$$\mathbb{E}\big[\exp(-\alpha T_{\text{couple}})\big] = \mathbb{E}\big[\exp(-\alpha T_2)\big].$$

We can express T_2 as the sum of (a) the Brownian hitting time T_1, being the time taken for B to pass from B_0 to $\frac{1}{2}(B_0 + \widetilde{B}_0)$, and (b) a randomized Brownian hitting time $T_2 - T_1$, being the time taken for B to pass from $\frac{1}{2}(B_0 + \widetilde{B}_0)$ to $M_1 = \sup\{B_t : t \leq T_1\}$. Then

$$T_2 = T_1 + H^1\big(\max\{S_0 \vee \widetilde{S}_0, M_1\} - \tfrac{1}{2}(B_0 + \widetilde{B}_0)\big),$$

where $H^1(a)$ is the time taken for a standard Brownian motion to pass from 0 to a.

We outline the calculations for the special case $B_0 = S_0$ and $\widetilde{B}_0 = \widetilde{S}_0$ (though the calculations can be extended to the general case). Thus, we are concerned with the moment generating function of $T_1 + H^1(M_1 - \frac{1}{2}(B_0 + \widetilde{B}_0))$.

The first task is to investigate aspects of the joint distribution of T_1 and M_1, specifically

$$Q_\alpha(a) = \mathbb{E}\big[\exp(-\alpha T_1); M_1 < a + B_0\big].$$

This can be calculated using excursion theory, for example by adapting the calculations of Rogers and Williams [34], §56. For convenience, we set $\alpha^* = \sqrt{2\alpha}$ and $b = B_0 - \frac{1}{2}(B_0 + \widetilde{B}_0) = \frac{1}{2}(B_0 - \widetilde{B}_0)$. Suppose that excursions are marked using an independent Poisson(α) point process on the time axis. Then we distinguish the following kinds of excursions of B from B_0, noting the rates at which they happen when the excursions are viewed as points of a Poisson process of excursions with respect to local time:

1. upward excursions which do not rise above the level $a + B_0$ but are marked (occurring at rate $\frac{1}{2}(\alpha^* \coth(\alpha^* a) - \frac{1}{a}))$;
2. upward excursions which rise above the level $a + B_0$ (occurring at rate $\frac{1}{2a}$);
3. downward excursions which do not fall below the level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ but are marked (occurring at rate $\frac{1}{2}(\alpha^* \coth(\alpha^* b) - \frac{1}{b}))$;
4. downward excursions which fall below the level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ (occurring at rate $\frac{1}{2b}$);
5. downward excursions which fall below the level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ but are not marked before hitting $\frac{1}{2}(B_0 + \widetilde{B}_0)$ (occurring at rate $\frac{\alpha^*}{2} \operatorname{cosech}(\alpha^* b)$).

These rates are computed as in the discussion of Rogers and Williams [34], Section 56, based on the identification of the law of the Brownian excursion discussed there. Thus, $Q_\alpha(a)$ can be computed as the probability that we see a downward excursion falling to level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ before it has been marked (that is, an excursion of type 5) *before* we ever see upward excursions rising to level $a + B_0$ (of type 2), or staying below this level but marked (type 1), or downward excursions which do not fall below the level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ but are marked (type 3), or downward excursions which fall below the level $\frac{1}{2}(B_0 + \widetilde{B}_0)$ but are marked before hitting $\frac{1}{2}(B_0 + \widetilde{B}_0)$ (type 4 but not type 5).

We can therefore use Poisson point process theory to compute

$$
\begin{aligned}
Q_\alpha(a) &= \frac{(\alpha^*/2)\operatorname{cosech}(\alpha^* b)}{1/(2a) + (1/2)(\alpha^* \coth(\alpha^* a) - 1/a) + (1/2)(\alpha^* \coth(\alpha^* b) - 1/b) + (1/(2b))} \\
&= \frac{\sinh(\alpha^* a)}{\sinh(\alpha^*(a + b))}.
\end{aligned}
\tag{9}
$$

Consequently, we can show that the desired moment generating function is given by

$$
\begin{aligned}
\mathbb{E}&\big[\exp(-\alpha T_{\text{couple}})\big] \\
&= \mathbb{E}\left[\exp\left(-\alpha\left(T_1 + H^1\left(M_1 - \frac{1}{2}(B_0 + \widetilde{B}_0)\right)\right)\right)\right] \\
&= \int_0^\infty \mathbb{E}\left[\exp\left(-\alpha H^1\left(a + B_0 - \frac{1}{2}(B_0 + \widetilde{B}_0)\right)\right)\right] Q_\alpha(\mathrm{d}a) \\
&= \int_0^\infty \mathbb{E}\big[\exp\big(-\alpha H^1(a + b)\big)\big] Q'_\alpha(a)\,\mathrm{d}a
\end{aligned}
$$

$$= \alpha^* \sinh(\alpha^* b) \int_0^\infty \frac{\mathbb{E}[\exp(-\alpha H^1(a+b))]}{\sinh^2(\alpha^*(a+b))} \, \mathrm{d}a$$

$$= \alpha^* \sinh(\alpha^* b) \int_0^\infty \frac{\mathrm{e}^{-\alpha^*(a+b)}}{\sinh^2(\alpha^*(a+b))} \, \mathrm{d}a = 1 + \sinh(\alpha^* b) \log \tanh\left(\frac{1}{2}\alpha^* b\right)$$

(using $b = \frac{1}{2}(B_0 - \widetilde{B}_0)$). □

These excursion-theoretic calculations have been checked by simulation, although direct simulation is computationally demanding because of the heavy tails of the Brownian hitting times involved in the reflection/synchronized coupling. This can to some extent be mitigated by the use of "Rao–Blackwellization" using the known formula for the moment generating function of the Brownian first passage time.

2.5. Comparison with maximal coupling

It is natural to ask whether the optimal immersed coupling is in fact a maximal coupling. Numerical evidence is that this is not the case, as is readily seen by computing the total mass of the minimum of the joint densities for (B_t, S_t) and $(\widetilde{B}_t, \widetilde{S}_t)$, and comparing with the coupling probability for optimal immersion coupling. From Revuz and Yor [32], Example 3.14 part 2°, the joint density of B_t and S_t (supposing $S_0 = B_0 = 0$) is given by

$$\sqrt{\frac{2}{\pi t}} \frac{2s-b}{t} \exp\left(-\frac{(2s-b)^2}{2t}\right) \qquad \text{for } b \leq s, s > 0. \tag{10}$$

(This follows readily from the reflection principle.) We can use this to evaluate numerically the moment generating function of the maximal coupling times (based on the maximal coupling derived from the method suggested by Sverchkov and Smirnov [38]). A numerical comparison of the moment generating functions for these couplings shows that they do not have the same distribution. This is numerical evidence that the optimal immersed coupling is distinct from (and slower than) the maximal coupling.

An explicit non-immersion maximal coupling for Brownian motion together with local time, for particular sets of randomized initial conditions, can be constructed by adapting the recipe of Pitman [31] for discrete time and space, and using the $2M - X$ theorem (Pitman [30]) to reduce the construction to the case of the scalar diffusion which is the three-dimensional Bessel process.

3. Equi-filtration couplings of Brownian motion together with local time

As noted in Section 2.2, the reflection/synchronized coupling is an equi-filtration coupling when viewed as a coupling for the absolute value $|X|$ of Brownian motion and $L^{(0)}$ its local time at 0. However the coupling is not equi-filtration when it is required to couple not just the absolute value $|X|$ but also X, the Brownian motion itself. This follows immediately from consideration

of the stochastic differential equation (5), which is of Tanaka type when viewed as generating the coupled Brownian motion $\widetilde{X}$ from $\int \text{sgn}(X) \, dX$. Unless X and $\widetilde{X}$ are identical, it is not possible to extract statistically appropriate signs for the excursions of $\widetilde{X}$ from the natural filtration $\{\mathcal{F}_t : t \geq 0\}$ of X.

The coupling of $(X, L^{(0)})$ with $(\widetilde{X}, \widetilde{L}^{(0)})$ can be modified to be equi-filtration by replacing $\text{sgn}(\widetilde{X}_t)$ and $\text{sgn}(X_t)$ in the coupling control by time-delayed versions, at the price of delaying the time of successful coupling. Given a positive continuous non-increasing function $\psi : [0, \infty) \to [0, \infty)$, we introduce the delayed time-change

$$\sigma(t) = t - \big(\psi(t) \wedge t\big). \tag{11}$$

For $t > 0$ we define a new coupled Brownian motion $\hat{X}_t$, starting at $\hat{X}_0 = \widetilde{X}_0$ but defined up to time T_2 as the solution to a time-delayed version of (5):

$$d\hat{X}_t = \text{sgn}(\hat{X}_{\sigma(t)}) J_t \, \text{sgn}(X_{\sigma(t)}) \, dX_t. \tag{12}$$

Here we use (4) to define the control J in terms of X *via* the stopping time T_1:

$$J_t = \begin{cases} -1 & \text{if } t \leq T_1, \\ +1 & \text{otherwise.} \end{cases}$$

Of course this exploits the remark in Section 2.2, that T_1 and T_2 can be defined as hitting times for $B = L^{(0)} - |X|$. Adopting the usual convention that $\text{sgn}(0) = 1$, and arguing from the positivity and continuity of ψ, we can solve the stochastic differential equation (12) step by step over successive small time intervals. This ensures that $\text{sgn}(\hat{X}_{\sigma(t)})$ and $\text{sgn}(X_{\sigma(t)})$ are defined and measurable with respect to $\mathcal{F}_{\sigma(t)} = \sigma\{X_s : s \leq \sigma(t)\}$.

This construction makes it clear that the coupled $\hat{X}$ is immersed in the natural filtration of X. However the reverse also holds:

Lemma 10. *For $\hat{X}$ defined in terms of X using the time-delayed stochastic differential equation* (12), *it is the case that X is immersed in the natural filtration of $\hat{X}$, so that the coupling of X and $\hat{X}$ is equi-filtration.*

Proof. The control J appearing in (12) satisfies $J \equiv -1$ up to the stopping time T_1. Accordingly $\hat{X} \equiv \hat{X}^*$ up to T_1, where

$$d\hat{X}_t^* = -\text{sgn}(\hat{X}_{\sigma(t)}^*) \, \text{sgn}(X_{\sigma(t)}) \, dX_t.$$

However, we may rewrite this last stochastic differential equation as

$$dX_t = -\text{sgn}(X_{\sigma(t)}) \, \text{sgn}(\hat{X}_{\sigma(t)}^*) \, d\hat{X}_t^*,$$

and so we find that

$$dX_t = -\text{sgn}(X_{\sigma(t)}) \, \text{sgn}(\hat{X}_{\sigma(t)}) \, d\hat{X}_t$$

holds up to time T_1. Since T_1 is a hitting time of X, it follows that X stopped at time T_1 is adapted to the filtration of $\hat{X}$, and thus that T_1 and thus J are adapted to the natural filtration of $\hat{X}$.

Arguing from time T_1 onwards, since $J_t = 1$ for $t > T_1$, we can re-write (12) as

$$\mathrm{d}X_t = \mathrm{sgn}(X_{\sigma(t)}) J_t \, \mathrm{sgn}(\hat{X}_{\sigma(t)}) \, \mathrm{d}\hat{X}_t = \mathrm{sgn}(X_{\sigma(t)}) \, \mathrm{sgn}(\hat{X}_{\sigma(t)}) \, \mathrm{d}\hat{X}_t \qquad \text{for } T_1 < t \le T_2.$$

It follows that X up to time T_2 is adapted to the natural filtration of $\hat{X}$. This establishes the mutual immersion property.   □

Of course $\hat{X} \neq \widetilde{X}$, and therefore we cannot assume that the equi-filtration coupling will have succeeded by time T_2. However, we can use the properties of the reflection/synchronized coupling to argue not only that the paths of $|\widetilde{X}|$ and $|\hat{X}|$ will be close in probability, but also that the same is true of the respective local times $\widetilde{L}^{(0)}$ and $\hat{L}^{(0)}$.

Lemma 11. *There is a sequence of equi-filtration couplings of $(\hat{X}^{(n)}, \hat{L}^{(0);(n)})$ with $(X, L^{(0)})$ such that $(|\hat{X}^{(n)}|, \hat{L}^{(0);(n)})$ converges in probability under supremum norm to $(|\widetilde{X}|, \widetilde{L}^{(0)})$* (*using the reflection/synchronized coupling*):

$$\mathbb{P}\left[\sup_t\left\{\left||\hat{X}_t^{(n)}| - |\widetilde{X}_t|\right| + \left|\hat{L}_t^{(0);(n)} - \widetilde{L}_t^{(0)}\right|\right\} > 4^{-n}\right] < 4^{-n}.$$

It suffices to exhibit a sequence for which the supremum converges to 0 in probability. Note that we do *not* assert that $\hat{X}^{(n)}$ converges to $\widetilde{X}$ in probability: if we were able to remove absolute values then this would contradict the known fact that the stochastic differential equation (5) of Tanaka type can have no strong solutions.

Proof of Lemma 11. To simplify notation, we take $X_0 = 0$. Note also that by construction $\hat{L}_0^{(0)} = \widetilde{L}_0^{(0)}$ and $\hat{X}_0 = \widetilde{X}_0$.

We shall select ψ depending on ε such that a suitably fast convergent sequence $\varepsilon_n \to 0$ delivers the required $(\hat{X}^{(n)}, \hat{L}^{(0);(n)})$.

First note that, if we set $\hat{S} = \hat{L}^{(0)}$ and $\hat{B} = \hat{L}^{(0)} - |\hat{X}|$, then

$$\mathrm{d}\hat{B} = \mathrm{sgn}(\hat{X}) \, \mathrm{sgn}(\hat{X}_\sigma) J \, \mathrm{sgn}(X_\sigma) \, \mathrm{d}X.$$

By choice of initial conditions, $\hat{B}_0 = \widetilde{B}_0$. It suffices to show convergence in probability of $\sup_t\{|\hat{B}_t - \widetilde{B}_t|\}$, the supremum norm of the difference. Because $|J| = |\mathrm{sgn}(X)| = |\mathrm{sgn}(\hat{X})| = 1$ we can use Doob's L^2 submartingale inequality, the L^2 isometry for Brownian stochastic integrals, and Jensen's inequality to deduce that

$$\mathbb{E}\left[\sup_t\left\{(\hat{B}_t - \widetilde{B}_t)^2\right\}\right]$$

$$= \mathbb{E}\left[\sup_t\left\{\left(\int_0^t \left(\mathrm{sgn}(\hat{X}) \, \mathrm{sgn}(\hat{X}_\sigma) J \, \mathrm{sgn}(X_\sigma) - J \, \mathrm{sgn}(X)\right) \mathrm{d}X\right)^2\right\}\right]$$

$$\le 4\mathbb{E}\left[\int_0^\infty \left|\mathrm{sgn}(\hat{X}_t) \, \mathrm{sgn}(\hat{X}_{\sigma(t)}) \, \mathrm{sgn}(X_{\sigma(t)}) - \mathrm{sgn}(X_t)\right|^2 \mathrm{d}t\right]$$

$$\leq 8 \int_0^\infty \mathbb{E}\big[\big|\mathrm{sgn}(\hat{X}_t)\,\mathrm{sgn}(\hat{X}_{\sigma(t)}) - 1\big|^2\big]\,\mathrm{d}t + 8 \int_0^\infty \mathbb{E}\big[\big|\mathrm{sgn}(X_{\sigma(t)}) - \mathrm{sgn}(X_t)\big|^2\big]\,\mathrm{d}t$$

$$= 32 \int_0^\infty \mathbb{P}\big[\mathrm{sgn}(\hat{X}_t) \neq \mathrm{sgn}(\hat{X}_{\sigma(t)})\big]\,\mathrm{d}t + 32 \int_0^\infty \mathbb{P}\big[\mathrm{sgn}(X_{\sigma(t)}) \neq \mathrm{sgn}(X_t)\big]\,\mathrm{d}t.$$

Both $\hat{X}$ and X are Brownian motions, though not necessarily begun at 0. It is therefore immediate that $\mathbb{P}[\mathrm{sgn}(\hat{X}_t) \neq \mathrm{sgn}(\hat{X}_{\sigma(t)})]$ and $\mathbb{P}[\mathrm{sgn}(X_{\sigma(t)}) \neq \mathrm{sgn}(X_t)]$ are both dominated by $\mathbb{P}[\mathrm{sgn}(\overline{X}_t) \neq \mathrm{sgn}(\overline{X}_{\sigma(t)})]$ for $\overline{X}$ a standard Brownian motion begun at 0 (since on average $\hat{X}$ will have to travel further to change sign than the standard Brownian motion $\overline{X}$). Moreover rotational symmetry of the standard bivariate normal distribution reveals, if $t > \psi(t)$,

$$\mathbb{P}\big[\mathrm{sgn}(\overline{X}_{\sigma(t)}) \neq \mathrm{sgn}(\overline{X}_t)\big] = \mathbb{P}\big[\mathrm{sgn}(\overline{X}_{\sigma(t)}) \neq \mathrm{sgn}\big(\overline{X}_{\sigma(t)} + (\overline{X}_t - \overline{X}_{\sigma(t)})\big)\big]$$

$$= \frac{1}{2}\mathbb{P}\big[|\overline{X}_{\sigma(t)}| < |\overline{X}_t - \overline{X}_{\sigma(t)}|\big]$$

$$= \frac{1}{2}\mathbb{P}\left[\frac{|\overline{X}_{\sigma(t)}|/\sqrt{t-\psi(t)}}{|(\overline{X}_t - \overline{X}_{\sigma(t)})|/\sqrt{\psi(t)}} < \sqrt{\frac{\psi(t)}{t-\psi(t)}}\right]$$

$$= \frac{1}{\pi}\tan^{-1}\sqrt{\frac{\psi(t)}{t-\psi(t)}}.$$

Thus we obtain, for $\varepsilon > \psi(\varepsilon)$,

$$\mathbb{E}\Big[\sup_t\{(\hat{B}_t - \widetilde{B}_t)^2\}\Big] \leq 64 \int_0^\infty \mathbb{P}\big[\mathrm{sgn}(\overline{X}_{\sigma(t)}) \neq \mathrm{sgn}(\overline{X}_t)\big]\,\mathrm{d}t$$

$$\leq 64\left(\varepsilon + \frac{1}{\pi}\int_\varepsilon^\infty \tan^{-1}\sqrt{\frac{\psi(t)}{t-\psi(t)}}\,\mathrm{d}t\right)$$

$$\leq 64\left(\varepsilon + \frac{1}{\pi}\int_\varepsilon^\infty \sqrt{\frac{\psi(t)}{t-\psi(t)}}\,\mathrm{d}t\right).$$

For any $\varepsilon \in (0, \frac{1}{2})$, we set

$$\psi(t) = \begin{cases} \varepsilon^3 & \text{if } t \in [0, \varepsilon), \\ \dfrac{\varepsilon^3}{(t-\varepsilon+1)^3} & \text{if } t \geq \varepsilon. \end{cases}$$

Then (a) ψ is positive continuous non-decreasing over $[0, \infty)$, (b) $\psi(t) \leq t$ for $t \geq \varepsilon$, and hence (c)

$$\mathbb{E}\Big[\sup_t\{(\hat{B}_t - \widetilde{B}_t)^2\}\Big] \leq 64\left(1 + \frac{1}{\pi}\int_\varepsilon^\infty \frac{1}{\sqrt{(t/\varepsilon)(t-\varepsilon+1)^3 - \varepsilon^2}}\,\mathrm{d}t\right)\varepsilon$$

$$\leq 64\left(1 + \frac{2}{\pi}\int_0^\infty \frac{1}{\sqrt{4(u+1)^3 - 1}}\,\mathrm{d}u\right)\varepsilon = 105.557\ldots \times \varepsilon.$$

The result follows, because L^2 convergence of random variables implies convergence in probability.   □

Theorem 12. *There are successful equi-filtration couplings of Brownian motion together with local time starting from any pair of initial conditions $(X_0, L_0^{(0)})$ and $(\widetilde{X}_0, \widetilde{L}_0^{(0)})$.*

The proof makes it plain that the coupling time for reflection/synchronized coupling can be approximated arbitrarily well in distribution by the coupling times for suitable equi-filtration couplings.

Proof of Theorem 12. It suffices to show that there is a sequence $\varepsilon_n \to 0$ such that the following holds for all reflection/synchronized couplings: if $||X_0| - |\widetilde{X}_0|| + |L_0^{(0)} - \widetilde{L}_0^{(0)}| < \varepsilon_n$ and $X_0 = 0$ then

$$\mathbb{P}\big[T_{\text{couple}} > 4^{-n}\big] \leq 4^{-n}. \tag{13}$$

For then we may construct a successful equi-filtration coupling as the concatenation of a series of equi-filtration couplings. In the first stage, Lemma 11 can be used to select an equi-filtration coupling producing $(\hat{X}, \hat{L}^{(0)})$, which approximates a reflection/synchronized coupling producing $(\widetilde{X}, \widetilde{L}^{(0)})$ (continued up to but not including time $T_2^{(1)}$, the end of the synchronized stage) such that we have the following control on the left-limits $\hat{X}_{T_2^{(1)}-}$, $\widetilde{X}_{T_2^{(1)}-}$, $\hat{L}_{T_2^{(1)}-}^{(0)}$ and $\widetilde{L}_{T_2^{(1)}-}^{(0)}$:

$$\mathbb{P}\big[||\hat{X}_{T_2^{(1)}-}| - |\widetilde{X}_{T_2^{(1)}-}|| + \big|\hat{L}_{T_2^{(1)}-}^{(0)} - \widetilde{L}_{T_2^{(1)}-}^{(0)}\big| > \varepsilon_1\big] < 4^{-1}.$$

Moreover, it follows from the construction of the reflection/synchronized coupling that $|X_{T_2^{(1)}-}| = |\widetilde{X}_{T_2^{(1)}-}| = 0$ while $L_{T_2^{(1)}-}^{(0)} = \widetilde{L}_{T_2^{(1)}-}^{(0)}$.

At time $T_2^{(1)}$, $\widetilde{X}$ makes a small jump so that $\widetilde{X}_{T_2^{(1)}} = \hat{X}_{T_2^{(1)}-}$, while X, $L^{(0)}$, $\widetilde{L}^{(0)}$, $\hat{X}$ and $\hat{L}^{(0)}$ trajectories remain continuous. The second and further stages are implemented by repeating the construction.

To be explicit, the construction continues through further stages $n = 2, 3, \ldots$, such that at the end $T_2^{(n)}-$ of stage n, conditional on successful fulfilment of all previous stages, we have $|X_{T_2^{(n)}-}| = |\widetilde{X}_{T_2^{(n)}-}| = 0$ and $L_{T_2^{(n)}-}^{(0)} = \widetilde{L}_{T_2^{(n)}-}^{(0)}$, moreover if $n \geq 2$ then $T_2^{(n)} - T_2^{(n-1)} < 4^{-n}$, and

$$\mathbb{P}\big[||\hat{X}_{T_2^{(n)}-}| - |\widetilde{X}_{T_2^{(n)}-}|| + \big|\hat{L}_{T_2^{(n)}-}^{(0)} - \widetilde{L}_{T_2^{(n)}-}^{(0)}\big| > \varepsilon_n\big] < 4^{-n}. \tag{14}$$

(We suppress the conditioning on previous stages for the sake of simple notation.) Stage n is implemented (a) by using a reflection/synchronized coupling of $(|\widetilde{X}|, \widetilde{L}^{(0)})$ with $(|X|, L^{(0)})$ which succeeds before time 4^{-n} (this has probability $1 - 4^{-n}$), and also (b) by invoking Lemma 11 to continue $(\hat{X}, \hat{L}^{(0)})$ by an equi-filtration coupling such that the maximum difference over this stage between the immersed coupling component $(|\widetilde{X}|, \widetilde{L}^{(0)})$ and the equi-filtration coupling

component $(|\hat{X}|, \hat{L}^{(0)})$ is less than ε_n (with conditional probability at least $1 - 4^{-n}$). It follows that the conditional probability of the nth stage ($n \geq 2$) completing successfully is at least $1 - 2 \times 4^{-n}$. At the end of stage n, we impose a small jump on $\widetilde{X}$ so that $\widetilde{X}_{T_2^{(n)}} = \hat{X}_{T_2^{(n)}-}$.

Thus, with total probability at least

$$1 - 4^{-1} - \sum_{n=2}^{\infty} 2 \times 4^{-n} = \frac{7}{12},$$

for all n, stage n is of duration less than 4^{-n}, and at time $T_2^{(n)}$ we have $X_{T_2^{(n)}-} = \widetilde{X}_{T_2^{(n)}-} = 0$, $L_{T_2^{(n)}-}^{(0)} = L_{T_2^{(n)}-}^{(0)}$, and $(X_{T_2^{(n)}-}, L_{T_2^{(n)}-}^{(0)})$ can be approximated by the equi-filtration coupling component $(\hat{X}_{T_2^{(n)}-}, \hat{L}_{T_2^{(n)}-}^{(0)})$ with sum of absolute differences less than ε_n.

It follows that the bound (13) can be used together with Lemma 11 to construct an equi-filtration coupling which has probability at least $\frac{7}{12}$ of succeeding in finite time. In case of default at any stage, one can then restart the sequence of couplings, so successful immersed coupling is almost sure to happen.

We now need to show that we can choose a sequence of ε_n to satisfy the bound (13). We refer to (B, S) coordinates. From $X_0 = 0$, we obtain $S_0 = B_0$ while $||X_0| - |\hat{X}_0|| = \hat{S}_0 - \hat{B}_0$ and $|L_0^{(0)} - \hat{L}_0^{(0)}| = |S_0 - \hat{S}_0|$. Suppose that $||X_0| - |\hat{X}_0|| < \varepsilon_n$ and $|L_0^{(0)} - \hat{L}_0^{(0)}| < \varepsilon_n$. Simple coupling arguments now show that the time taken to achieve reflection/synchronized coupling from such starting points is maximized if

$$L_0^{(0)} - \hat{L}_0^{(0)} = S_0 - \hat{S}_0 = \varepsilon_n,$$
$$||X_0| - |\hat{X}_0|| = \hat{S}_0 - \hat{B}_0 = \varepsilon_n.$$

But with such initial conditions, we can apply a scaling argument to show that T_2/ε^2 has a distribution not depending on ε_n. It therefore follows that we can choose ε_n to ensure that (13) holds whenever $||X_0| - |\widetilde{X}_0|| + |L_0^{(0)} - \widetilde{L}_0^{(0)}| < \varepsilon_n$ and $X_0 = 0$. $\qquad \square$

The following corollary will be of assistance when constructing equi-filtration couplings of BKR diffusions in the next section.

Corollary 13. *The equi-filtration coupling of Theorem 12 can be localized in the following sense*: *for fixed $\delta > 0$, for all sufficiently small $\varepsilon > 0$ and for all pairs of initial conditions $(X_0, L_0^{(0)})$ and $(\widetilde{X}_0, \widetilde{L}_0^{(0)})$ with*

$$\left||X_0| - |\widetilde{X}_0|\right| + \left|L_0^{(0)} - \widetilde{L}_0^{(0)}\right| < \varepsilon,$$

we can construct a successful equi-filtration coupling of $(X, L^{(0)})$ and $(\widetilde{X}, \widetilde{L}^{(0)})$ such that

$$\mathbb{P}\left[\text{one of } (X, L^{(0)}) \text{ and } (\widetilde{X}, \widetilde{L}^{(0)}) \text{ does not stay within } \mathrm{ball}\left((X_0, L_0^{(0)}), \delta\right)\right] < \varepsilon.$$

Proof. From the proof of Theorem 12, it follows that for sufficiently small ε we can obtain uniformly arbitrarily small probability of the equi-filtration coupling failing to couple within an arbitrarily small period of time. The corollary then follows by observing that it follows from continuity of Brownian motion and Brownian local time that over a sufficiently small period of time we can ensure that the probability of there being large deviations either of motion or of local time is arbitrarily small. □

4. Application to BKR diffusions

The BKR diffusion was introduced by Beneš, Karatzas and Rishel [4] as part of an investigation into a control problem which possesses no strict-sense optimal law: it is a two-dimensional diffusion (X, Y) for which X and Y are Brownian motions connected by

$$\text{sgn}(X)\,\mathrm{d}X + \text{sgn}(Y)\,\mathrm{d}Y = 0. \tag{15}$$

Essentially (X, Y) diffuses linearly in a Brownian fashion along the boundary of a square $\{(x, y): |x| + |y| = \ell\}$, except that ℓ increases according to a local time driving term whenever (X, Y) visits one of the vertices of the square. Émery [13] studied filtration questions concerning this process, and in particular showed that the natural filtration of (X, Y) is Brownian even when (X, Y) is started from the origin. In this connection, he asked a specific question (Émery [13], Question), which can be stated concisely as follows: given two initial points (X_0, Y_0) and $(\widetilde{X}_0, \widetilde{Y}_0)$, neither of which is the origin, is there an almost surely successful equi-filtration coupling of BKR diffusions (X, Y) and $(\widetilde{X}, \widetilde{Y})$ started from these two points? An affirmative answer would lead to a constructive proof of Brownianity of the filtration of (X, Y) started from the origin. An immersed coupling is exhibited in Émery [13], Lemma 5; however equi-filtration is required for the purposes of obtaining a constructive proof of the filtration result.

4.1. Sketch construction of immersed coupling for the BKR diffusion

The work of Sections 2, 3 suggests a direct strategy for constructing successful equi-filtration couplings of BKR diffusions; start with an almost-surely successful immersed coupling, and then perturb it to produce a nearly-successful coupling which can then be iterated to generate a successful coupling following the methods used to prove Theorem 12 and associated lemmas. The immersed coupling described in Émery [13], Lemma 5, bears a strong family resemblance to the reflection/synchronized coupling given above (Example 3): in preparation for construction of the equi-filtration coupling we first sketch a description of Émery's immersed coupling for BKR diffusions using the terminology of our paper.

Application of the Tanaka formula for Brownian local time to (15) shows that $|X| + |Y| = L^{X;(0)} + L^{Y;(0)}$ increases as the sum of the local times accumulated by X and Y at 0. In the case when (X, Y) does not begin at the origin, so that $h = \frac{1}{2}(|X_0| + |Y_0|) > 0$, we can determine a real Brownian motion which drives the BKR diffusion by first defining a binary càdlàg switch process K, which takes values 0 or 1, changing only according to the following rules:
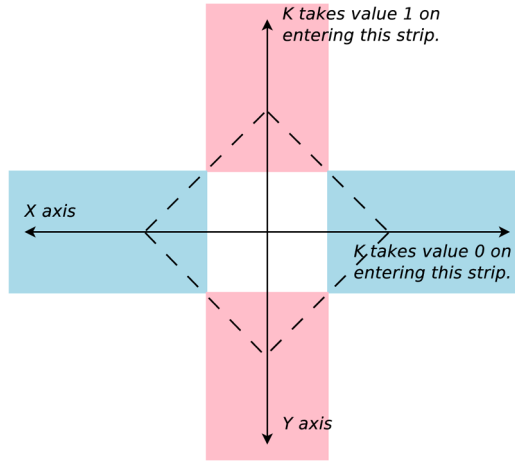
Figure 2. The construction of a BKR diffusion (X, Y) from a real Brownian motion can be based on a binary càdlàg process K, which switches when it enters specific regions with boundaries determined by $x, y = \pm h$ for $h = \frac{1}{2}(|X_0| + |Y_0|)$. The initial value (X_0, Y_0) lies on the boundary of the diamond-shaped region, and by construction the diffusion will never enter the interior of this region.

1. K takes value 1 on entry to the region $|X| < h$;
2. K takes value 0 on entry to the region $|Y| < h$;

and otherwise K is time-constant. The initial value of K is given by

$$K_0 = \begin{cases} 1 & \text{if } |X_0| \leq h, \\ 0 & \text{otherwise.} \end{cases}$$

This construction is illustrated in Figure 2.

We can then define a real-valued Brownian motion A by

$$dA = K \operatorname{sgn}(Y) dX - (1 - K) \operatorname{sgn}(X) dY. \tag{16}$$

Note that the definition of K implies that Y never vanishes when $K = 1$, and X never vanishes when $K = 0$. This allows us to use A to construct (X, Y) as follows:

$$dX = K \operatorname{sgn}(Y) dA - (1 - K) \operatorname{sgn}(X) \operatorname{sgn}(Y) dY,$$
$$dY = -K \operatorname{sgn}(X) \operatorname{sgn}(Y) dX - (1 - K) \operatorname{sgn}(X) dA. \tag{17}$$

Thus if $K = 1$, then we can construct X in terms of A (since Y does not then change sign) and then Y in terms of X, and conversely if $K = 0$, then we can construct Y in terms of A and then X in terms of Y. In particular, it is convenient to note the following relationships between X, Y and the driving Brownian motion A:

$$dX = \operatorname{sgn}(Y) dA, \tag{18}$$

$$dY = -\operatorname{sgn}(X) dA. \tag{19}$$

We can now sketch the construction of an immersed coupling with another BKR diffusion which also does not begin at the origin. Let $(\widetilde{X}, \widetilde{Y})$ be the coupled BKR diffusion, based on $\widetilde{h} = \frac{1}{2}(|\widetilde{X}_0| + |\widetilde{Y}_0|) > 0$, with $\widetilde{K}$ and $\widetilde{A}$ defined in direct analogy to K and A. We define the coupling (a variant of the reflection/synchronized coupling described in Definition 3) by the following definition.

Definition 14 (Variant reflection/synchronized coupling). *With the above notation, two BKR diffusions (X, Y) and $(\widetilde{X}, \widetilde{Y})$ (adapted to the same filtration) are said to be in* reflection/synchronized coupling *if their driving Brownian motions A and $\widetilde{A}$ are related by*

$$\mathrm{d}\widetilde{A} = \mathrm{sgn}(\widetilde{Y}) J \, \mathrm{sgn}(Y) \, \mathrm{d}A, \tag{20}$$

where $J = -1$ until X and $\widetilde{X}$ first meet, after which we set $J = +1$. (Thus $\widetilde{X}$ is a reflection of X till $X = \widetilde{X}$.)

Note that this variant coupling does not treat X and Y symmetrically. As discussed in detail by Émery [13], the variant reflection/synchronized coupling can be represented as an immersed coupling and is almost surely successful. Given (X, Y), in order to reconstruct $(\widetilde{X}, \widetilde{Y})$ it is necessary to augment the filtration using an appropriate independent sequence of equiprobable ± 1 random variables; this is because it is not possible to obtain strong solutions to the stochastic differential system given by (20), (17), and the analogue of (17) giving $(\widetilde{X}, \widetilde{Y})$ in terms of $\widetilde{A}$. Note however that we *can* construct the paths of $\widetilde{X}$ and $|\widetilde{Y}|$ from the path of (X, Y): indeed

Lemma 15. *If (X, Y) and $(\widetilde{X}, \widetilde{Y})$ are connected by the variant reflection/synchronized coupling specified in Definition 14 then $\widetilde{X}$ and $|\widetilde{Y}|$ are both adapted to the natural filtration of X, and $T_1 = \inf\{t\colon X_t = \widetilde{X}_t\}$, and the coupling time $T_{\mathrm{couple}} = T_2 = \inf\{t > T_1\colon |Y_t| = |\widetilde{Y}_t| = 0\}$ are stopping times for this filtration. In particular T_{couple} is almost surely finite.*

Proof. For the case of $\widetilde{X}$, observe that T_1 is the first time t that $\widetilde{X}_0 - (X_t - X_0)$ hits $\frac{1}{2}(X_0 + \widetilde{X}_0)$, and therefore T_1 is a stopping time for the natural filtration of X. Moreover,

$$\widetilde{X} = \left(\widetilde{X}_0 - (X_{t \wedge T_1} - X_0)\right) + (X_{t \vee T_1} - X_{T_1});$$

hence $\widetilde{X}$ is also adapted to the natural filtration of X.

For the case of $|\widetilde{Y}|$, observe that $|\widetilde{Y}|$ satisfies

$$\mathrm{d}|\widetilde{Y}| = \mathrm{sgn}(\widetilde{Y}) \, \mathrm{d}\widetilde{Y} + \mathrm{d}L^{\widetilde{Y};(0)},$$

where $L^{\widetilde{Y};(0)}$ is the local time accumulated by $\widetilde{Y}$ at 0, and we can use the Lévy transform to show that it suffices to establish that $\int_0^t \mathrm{sgn}(\widetilde{Y}) \, \mathrm{d}\widetilde{Y}$ is measurable with respect to the natural filtration of X. But we can employ the stochastic differential equation (20) determining the coupling, together with (19);

$$\mathrm{sgn}(\widetilde{Y}) \, \mathrm{d}\widetilde{Y} = -\mathrm{sgn}(\widetilde{X}) \, \mathrm{sgn}(\widetilde{Y}) \, \mathrm{d}\widetilde{A} = -\mathrm{sgn}(\widetilde{X}) \, \mathrm{sgn}(\widetilde{Y}) \, \mathrm{sgn}(\widetilde{Y}) J \, \mathrm{sgn}(Y) \, \mathrm{d}A$$
$$= -\mathrm{sgn}(\widetilde{X}) J \, \mathrm{sgn}(Y) \, \mathrm{d}A.$$

Thus, we can deduce that $\int_0^t \text{sgn}(\widetilde{Y}) \, d\widetilde{Y}$ is given by a Brownian stochastic integral adapted to the natural filtration of X, since $dX = \text{sgn}(Y) \, dA$ by (18), and we have already shown that $\widetilde{X}$ is so adapted.

Moreover after T_1 we have $J \equiv 1$ and $X \equiv \widetilde{X}$, hence

$$
\begin{aligned}
d|\widetilde{Y}| - d|Y| &= \text{sgn}(\widetilde{Y}) \, d\widetilde{Y} - \text{sgn}(Y) \, dY + dL^{\widetilde{Y}:(0)} - dL^{Y:(0)} \\
&= -\text{sgn}(\widetilde{X}) \, \text{sgn}(Y) \, dA - \text{sgn}(Y) \, dY + dL^{\widetilde{Y}:(0)} - dL^{Y:(0)} \\
&= -\text{sgn}(X) \, \text{sgn}(Y) \, dA - \text{sgn}(Y) \, dY + dL^{\widetilde{Y}:(0)} - dL^{Y:(0)} = dL^{\widetilde{Y}:(0)} - dL^{Y:(0)},
\end{aligned}
$$

where we use (19) to cancel the two Brownian stochastic differentials. Thus after T_1 it is the case that either $|Y| \geq |\widetilde{Y}|$ for all time, or $|Y| \leq |\widetilde{Y}|$ for all time, and coupling occurs at the first time T_2 after T_1 that the coupled reflected Brownian motions $|Y|$ and $|\widetilde{Y}|$ simultaneously hit 0.

Now $|Y|$ hits zero when $|X|$ first hits $|X_0| + |Y_0|$, or at subsequent times when $|X|$ attains its running supremum, while we have already shown that $|\widetilde{Y}|$ is adapted to the natural filtration of X, therefore T_2 is a stopping time for this filtration.

Finally, almost sure finiteness of T_{couple} follows, since it is the (dependent) sum of two almost surely finite Brownian hitting times T_1 and $T_2 - T_1$. □

We shall use this partial reconstruction when analyzing the equi-filtration coupling described below.

4.2. An equi-filtration coupling for BKR diffusions

In order to construct an equi-filtration coupling for BKR diffusions (X, Y) and $(\widetilde{X}, \widetilde{Y})$, with neither BKR diffusion starting at the origin, we adopt the strategy of Section 3. Given a delayed time-change $\sigma(t)$ defined in terms of a positive continuous non-increasing function ψ as in (11), we can define a new driving Brownian motion $\hat{A}$ in terms of A via

$$
d\hat{A}_t = \text{sgn}(\hat{Y}_{\sigma(t)}) J_t \, \text{sgn}(Y_{\sigma(t)}) \, dA. \tag{21}
$$

Here X (and Y) are defined in terms of A using (17); we note that J is the *immersed* control given in the previous subsection, constructed in terms of (X, Y) by setting $J = -1$ till the time T_1 when X first hits $\frac{1}{2}(X_0 + \widetilde{X}_0)$, and then setting $J = +1$; finally, $\hat{Y}$ (and $\hat{X}$) are defined in terms of $\hat{A}$ using the analogue of (17). The use of the delay σ means that the system of these stochastic differential equations has a unique strong solution so long as neither BKR diffusion is begun at the origin. Note that there are issues in finding strong solutions to (17) together with the switching processes K and $\hat{K}$ if either or both of the BKR diffusions start at the origin.

Lemma 16. *Suppose $(\hat{X}, \hat{Y})$ is a BKR diffusion defined in terms of a BKR diffusion (X, Y) using the time-delayed stochastic differential equation* (21) *and the analogues of the defining equation*

(17) *together with switching processes* K *and* $\hat{K}$, *so that*

$$d\hat{X} = \operatorname{sgn}(\hat{Y}) \, d\hat{A}, \tag{22}$$

$$d\hat{Y} = -\operatorname{sgn}(\hat{X}) \, d\hat{A}. \tag{23}$$

If neither BKR diffusion is begun at the origin, then the resulting coupling is equi-filtration.

Note that this definition is not symmetrical in (X, Y) and $(\hat{X}, \hat{Y})$, since the coupling control J is defined in terms of (X, Y). Note further that we do not assert that the coupling is successful!

Proof of Lemma 16. It follows from construction that $(\hat{X}, \hat{Y})$ is immersed in the filtration of (X, Y). On the other hand we can argue as in Lemma 10 that the reverse also holds, and hence that this coupling is equi-filtration. As in Lemma 10 of the argument for the case of Brownian motion with local time, the key point is to argue first that the trajectory of (X, Y) up to the time T_1 (while $J = -1$) is immersed in the filtration of $(\hat{X}, \hat{Y})$, and then to argue that the subsequent construction (while $J = +1$) is also immersed. The crucial point is that T_1 is a stopping time for the filtration of $(\hat{X}, \hat{Y})$, as noted in Lemma 15. $\square$

Because of Lemma 15, it makes sense to discuss $\widetilde{X}_{T_2}$ and $|\widetilde{Y}_{T_2}|$ defined in terms of X and Y, and in particular to consider the extent to which $\hat{X}_{T_2}$ and $|\hat{Y}_{T_2}|$ differ from $\widetilde{X}_{T_2}$ and $|\widetilde{Y}_{T_2}|$. Moreover $|\widetilde{Y}|_{T_2} = Y_{T_2} = 0$, so control of $||\hat{Y}_{T_2}| - |\widetilde{Y}_{T_2}||$ corresponds directly to control of $|\hat{Y}_{T_2} - \widetilde{Y}_{T_2}| = |\hat{Y}_{T_2}|$ itself.

Lemma 17. *Suppose* $(\hat{X}, \hat{Y})$ *is a BKR diffusion defined in terms of a BKR diffusion* (X, Y) *using the time-delayed stochastic differential equation* (21) *and the analogues of the defining equations* (18) *and* (19). *For any* $\delta > 0$, *we can choose* $\varepsilon \in (0, \frac{1}{2})$ *sufficiently small so that if the time-delay* $\sigma(t) = t - (\psi(t) \wedge t)$ *is defined via* $\psi(t) = \varepsilon^3 / ((t - \varepsilon + 1)^3)$ *for* $t \geq \varepsilon$ *then*

$$\mathbb{P}\big[|\hat{X}_{T_2} - \widetilde{X}_{T_2}| + |\hat{Y}_{T_2} - \widetilde{Y}_{T_2}| > \delta\big] \leq \delta. \tag{24}$$

Proof. Consider the stochastic differential equation for $\hat{X}$:

$$d\hat{X} = \operatorname{sgn}(\hat{Y}) \, d\hat{A} = J \operatorname{sgn}(Y_\sigma) \operatorname{sgn}(\hat{Y}_\sigma) \operatorname{sgn}(\hat{Y}) \, dA.$$

Since $dX = \operatorname{sgn}(Y) \, dA$, and since $Y \approx Y_\sigma$ and $\hat{Y} \approx \hat{Y}_\sigma$, it follows that the coupling between $\hat{X}$ and X approximates a reflection coupling up to time T_1, and after that approximates a synchronized coupling. Calculating as in Lemma 11, but recalling from Lemma 15 that $\widetilde{X} = (\widetilde{X}_0 - (X_{t \wedge T_1} - X_0)) + (X_{t \vee T_1} - X_{T_1})$ is actually adapted to the filtration of (X, Y),

$$\mathbb{E}\Big[\sup_t\big\{(\hat{X}_t - \widetilde{X}_t)^2\big\}\Big] \leq 4 \int_0^\infty \mathbb{E}\big[\big(J \operatorname{sgn}(Y_\sigma) \operatorname{sgn}(\hat{Y}_\sigma) \operatorname{sgn}(\hat{Y}) - J \operatorname{sgn}(Y)\big)^2\big] \, dt$$

$$\leq 32 \int_0^\infty \mathbb{P}\big[\operatorname{sgn}(\hat{Y}_\sigma) \neq \operatorname{sgn}(\hat{Y})\big] \, dt + 32 \int_0^\infty \mathbb{P}\big[\operatorname{sgn}(Y_\sigma) \neq \operatorname{sgn}(Y)\big] \, dt$$

$$\leq 105.557 \ldots \times \varepsilon.$$

Thus we can control the extent to which the approximate reflection coupling of X and $\hat{X}$ will deviate from the reflection coupling of X and $\widetilde{X}$ by using the Markov inequality: for any $\delta > 0$ such that $\varepsilon < \delta^3/(12 \times 105.557\ldots)$, it follows that

$$\mathbb{P}\left[\sup_t \{|\hat{X}_t - \widetilde{X}_t|\} > \delta/2\right] \leq \delta/3. \tag{25}$$

We now need to control the approximation of $|\widetilde{Y}|$ by $|\hat{Y}|$. We first use the almost-sure finiteness of the stopping time T_2 to select a constant time $t_{\max}$ such that

$$\mathbb{P}[T_2 > t_{\max}] < \delta/3. \tag{26}$$

It suffices to control the approximation of $\int_0^t \mathrm{sgn}(\widetilde{Y})\,\mathrm{d}\widetilde{Y}$ by $\int_0^t \mathrm{sgn}(\hat{Y})\,\mathrm{d}\hat{Y}$ when $0 \leq t \leq t_{\max}$. Observe that

$$\mathrm{sgn}(\hat{Y})\,\mathrm{d}\hat{Y} - \mathrm{sgn}(\widetilde{Y})\,\mathrm{d}\widetilde{Y} = -\mathrm{sgn}(\hat{Y})\,\mathrm{sgn}(\hat{X})\,\mathrm{d}\hat{A} + \mathrm{sgn}(\widetilde{Y})\,\mathrm{sgn}(\widetilde{X})\,\mathrm{d}\widetilde{A}$$

$$= -J\,\mathrm{sgn}(\hat{Y})\,\mathrm{sgn}(\hat{X})\,\mathrm{sgn}(\hat{Y}_\sigma)\,\mathrm{sgn}(Y_\sigma)\,\mathrm{d}A + J\,\mathrm{sgn}(Y)\,\mathrm{sgn}(\widetilde{X})\,\mathrm{d}A.$$

Hence, (using the same definition of ψ)

$$\mathbb{E}\left[\sup_{t \leq t_{\max}}\left\{\left(\int_0^t \mathrm{sgn}(\hat{Y})\,\mathrm{d}\hat{Y} - \int_0^t \mathrm{sgn}(\widetilde{Y})\,\mathrm{d}\widetilde{Y}\right)^2\right\}\right]$$

$$\leq 4\int_0^{t_{\max}} \mathbb{E}\left[\left(\mathrm{sgn}(\hat{Y})\,\mathrm{sgn}(\hat{X})\,\mathrm{sgn}(\hat{Y}_\sigma)\,\mathrm{sgn}(Y_\sigma) - \mathrm{sgn}(Y)\,\mathrm{sgn}(\widetilde{X})\right)^2\right]\mathrm{d}t$$

$$\leq 48\int_0^\infty \mathbb{P}\left[\mathrm{sgn}(\hat{Y}) \neq \mathrm{sgn}(\hat{Y}_\sigma)\right]\mathrm{d}t + 48\int_0^\infty \mathbb{P}\left[\mathrm{sgn}(Y) \neq \mathrm{sgn}(Y_\sigma)\right]\mathrm{d}t$$

$$+ 48\int_0^{t_{\max}} \mathbb{P}\left[\mathrm{sgn}(\hat{X}) \neq \mathrm{sgn}(\widetilde{X})\right]\mathrm{d}t$$

$$\leq 158.336\ldots \times \varepsilon + 48\int_0^{t_{\max}} \mathbb{P}\left[\mathrm{sgn}(\hat{X}) \neq \mathrm{sgn}(\widetilde{X})\right]\mathrm{d}t.$$

Now

$$\int_0^{t_{\max}} \mathbb{P}\left[\mathrm{sgn}(\hat{X}) \neq \mathrm{sgn}(\widetilde{X})\right]\mathrm{d}t \leq \int_0^{t_{\max}} \left(\mathbb{P}\left[|\hat{X} - \widetilde{X}| \geq \eta\right] + \mathbb{P}\left[|\widetilde{X}| < \eta\right]\right)\mathrm{d}t$$

$$\leq \int_0^{t_{\max}} \left(\frac{1}{\eta^2}\mathbb{E}\left[|\hat{X} - \widetilde{X}|^2\right] + 1 \wedge \frac{2\eta}{\sqrt{2\pi t}}\right)\mathrm{d}t$$

$$\leq \frac{105.557\ldots \times t_{\max}}{\eta^2}\varepsilon + \frac{2}{\pi}\eta^2 + \sqrt{\frac{8t_{\max}}{\pi}}\eta.$$

It follows that if we choose first η then ε small enough that

$$\eta < \frac{1}{12 \times 48} \sqrt{\sqrt{\frac{\pi}{8t_{\max}}} \frac{\delta^3}{12}},$$

$$\eta < \sqrt{\frac{\pi}{96 \times 12} \frac{\delta^3}{12}},$$

$$\varepsilon < \frac{\eta^2}{12 \times 48 \times 105.557\ldots \times t_{\max}} \frac{\delta^3}{12},$$

$$\varepsilon < \frac{1}{12 \times 158.336\ldots} \frac{\delta^3}{12},$$

then

$$\mathbb{E}\left[\sup_{t \le t_{\max}} \left\{ \left(\int_0^t \operatorname{sgn}(\hat{Y})\,\mathrm{d}\hat{Y} - \int_0^t \operatorname{sgn}(\widetilde{Y})\,\mathrm{d}\widetilde{Y} \right)^2 \right\}\right]$$

$$\le 158.336\ldots \times \varepsilon + 48\left(\frac{105.557\ldots \times t_{\max}}{\eta^2}\varepsilon + \frac{2}{\pi}\eta^2 + \sqrt{\frac{8t_{\max}}{\pi}}\eta \right) \le \frac{1}{3} \times \frac{\delta^3}{12}. \tag{27}$$

The lemma now follows from the inequalities (25), (26), and an application of the Markov inequality to (27). □

We can now state and prove the main result of this section, that BKR diffusions begun at non-zero points can be coupled in an equi-filtration manner.

Theorem 18. *Given two BKR diffusions begun at different non-zero initial points, they can be coupled in a mutually immersed manner which succeeds in almost surely finite time, using an infinite sequence of increasingly rapid equi-filtration couplings, each of which approximates the variant reflection/synchronized coupling but with delays built into the reflection and synchronization couplings to render them equi-filtration.*

Proof. From Lemma 17 it follows that at time T_2 we can bring (X, Y) and $(\hat{X}, \hat{Y})$ arbitrarily close together with probability arbitrarily close to 1. Moreover $Y_{T_2} = \widetilde{Y}_{T_2} = 0$, so that $\hat{Y}_{T_2}$ will be arbitrarily close to zero with probability arbitrarily close to 1.

Restarting at time T_2, the evolutions of $(X, |Y|)$ and $(\hat{X}, |\hat{Y}|)$ can be related to the behaviour of Brownian motion and its local time at 0. Specifically, $(|Y|, X - |Y|)$ (respectively $(|\hat{Y}|, \hat{X} - |\widetilde{Y}|)$) has the stochastic dynamics of the absolute value of Brownian motion together with its local time at 0, at least until X (respectively $\hat{X}$) hits zero.

We can therefore apply the iterative coupling technique of Section 3 to achieve exact coupling of (X, Y) and $(\hat{X}, \hat{Y})$; the localization supplied by Corollary 13 implies that there is a positive probability of achieving this coupling before either X or $\hat{X}$ hit 0, with a uniform positive lower bound on the probability which tends to 1 as the restarted values at T_2 of $X - \hat{X}$, Y and $\hat{Y}$ tend to zero. In the event of default (i.e., the initial variant reflection/synchronized coupling fails

to achieve approximate coupling at T_2, or X or $\hat{X}$ hits 0 subsequent to T_2), then the whole coupling procedure can be restarted; the initial delayed variant reflection/synchronized coupling can be arranged to deliver approximate coupling to an arbitrarily small tolerance with probability arbitrarily close to 1, and then the subsequent iterative coupling will also have success probability arbitrarily close to 1. Thus, it is possible to arrange that the coupling procedure will only need to be restarted a finite number of times before coupling is achieved. $\square$

5. Conclusion

In this paper, we have established the basic properties of the reflection/synchronized coupling for Brownian motion together with local time, and in particular we have shown it is an optimal immersed. but not maximal, coupling. We have also shown that this coupling can be approximated by explicit equi-filtration couplings; moreover that it can be used as the basis for immersion and equi-filtration couplings of a more complicated diffusion, thus answering a question arising from filtration theory.

Further questions include the following:

1. Is Émery's [13] immersed BKR coupling optimal amongst all immersed couplings of BKR diffusions? We would be surprised if this were the case, since there are two distinct variants of Émery's [13] coupling depending on whether the construction is based on the X component or the Y component: it seems implausible that optimal couplings would permit such a symmetry. It is of course possible that there is *no* optimal immersed coupling: it may be the case that some strategies work better for short time while others work better for long time.

2. Reverting to the case of Brownian motion with local time, is it possible to couple Brownian motion together with local times accumulated at two or more distinct points? This seems to be a hard question. Analogous generalizations have been carried out for the case of coupling Brownian motions together with iterated time-integrals (Kendall and Price [23]); however there is a useful linearity in the time-integral case which is not present here. The question of coupling finite sets of local times could be viewed as a question of whether one could couple a finite version of the celebrated Brownian burglar (Warren and Yor [41], Aldous [1]).

3. There is interesting territory to be explored in the realm of couplings which fall short of being maximal but yet are not immersed. One example in applied stochastic process theory is supplied by Smith [37], who investigates the mixing time of a simple Gibbs sampler on the unit simplex using a two-stage coupling of which the first is immersed (Markovian, in Smith's chosen terminology) while the second couples an associated partition process anticipatively. This non-immersed coupling allows Smith to give an affirmative answer to a conjecture by Aldous concerning the mixing time of this Gibbs sampler. Arguably Sigman's [36] perfect simulation algorithm for super-stable $M/G/c$ queues can be put in the same category, as this depends on coupling service times not according to time of arrival of customer but according to time of start of service. It would be likely to be most illuminating if one could discover simple Brownian coupling problems for which gains of a similar kind can be made.

4. It seems clear that the techniques of this paper can be generalized to show that immersion couplings of suitably regular diffusions can always be approximated by equi-filtration couplings, and it would be interesting to see a fully rigorous proof in a case where the qualification "suitably regular" is given a pleasant and natural meaning.

The underlying reflection/synchronization coupling for Brownian motion together with local time is extremely simple, and lends itself to rather complete calculation. Not only is it an example of the general programme of coupling Brownian motion together with appropriate functionals (Ben Arous, Cranston, and Kendall [3], Kendall and Price [23], Kendall [20,22]), but also it can be viewed as a basic coupling strategy that, like the reflection coupling of Brownian motion (Lindvall [26]) has the potential to serve as a model in much more general situations. The application to the BKR diffusion in this paper illustrates this point; it is hoped that the calculations described above will facilitate the use of the reflection/synchronization coupling as a building block in other applications of coupling to probability theory.

Acknowledgements

References

[1] Aldous, D.J. (1998). Brownian excursion conditioned on its local time. *Electron. Commun. Probab.* **3** 79–90 (electronic). MR1650567

[2] Beghdadi-Sakrani, S. and Émery, M. (1999). On certain probabilities equivalent to coin-tossing, d'après Schachermayer. In *Séminaire de Probabilités*, *XXXIII. Lecture Notes in Math.* **1709** 240–256. Berlin: Springer. MR1767998

[3] Ben Arous, G., Cranston, M. and Kendall, W.S. (1995). Coupling constructions for hypoelliptic diffusions: Two examples. In *Stochastic Analysis* (*Ithaca*, *NY*, 1993) (M. Cranston and M. Pinsky, eds.). *Proc. Sympos. Pure Math.* **57** 193–212. Providence, RI: Amer. Math. Soc. MR1335472

[4] Beneš, V.E., Karatzas, I. and Rishel, R.W. (1991). The separation principle for a Bayesian adaptive control problem with no strict-sense optimal law. In *Applied Stochastic Analysis* (*London*, 1989). *Stochastics Monogr.* **5** 121–156. New York: Gordon & Breach. MR1108420

[5] Burdzy, K. and Kendall, W.S. (2000). Efficient Markovian couplings: Examples and counterexamples. *Ann. Appl. Probab.* **10** 362–409. MR1768241

[6] Chen, M.-F. (1994). Optimal couplings and application to Riemannian geometry. In *Probability Theory and Mathematical Statistics* (*Vilnius*, 1993) 121–142. Vilnius: TEV. MR1649575

[7] Chen, M.F. and Li, S.F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.* **17** 151–177. MR0972776

[8] Connor, S. (2013). Optimal coadapted coupling for a random walk on the hyper-complete graph. *J. Appl. Probab.* **50** 1117–1130. MR3161377

[9] Connor, S. and Jacka, S. (2008). Optimal co-adapted coupling for the symmetric random walk on the hypercube. *J. Appl. Probab.* **45** 703–713. MR2455179

[10] Connor, S.B. (2007). Coupling: Cutoffs, CFTP and tameness. Ph.D. Thesis, Univ. Warwick.

[11] Doeblin, W. (1938). Exposé de la Théorie des Chaînes simples constants de Markoff á un nombre fini d'États. *Revue Math. de l'Union Interbalkanique* **2** 77–105.

[12] Émery, M. (2005). On certain almost Brownian filtrations. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 285–305. MR2139021

[13] Émery, M. (2009). Recognising whether a filtration is Brownian: A case study. In *Séminaire de Probabilités XLII. Lecture Notes in Math.* **1979** 383–396. Berlin: Springer. MR2599217

[14] Goldstein, S. (1978/79). Maximal coupling. *Z. Wahrsch. Verw. Gebiete* **46** 193–204. MR0516740

[15] Griffeath, D. (1974/75). A maximal coupling for Markov chains. *Z. Wahrsch. Verw. Gebiete* **31** 95–106. MR0370771

[16] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. Amsterdam: North-Holland. MR1011252

[17] Itô, K. (1974/75). Stochastic differentials. *Appl. Math. Optim.* **1** 374–381. MR0388538

[18] Kendall, W.S. (1998). From stochastic parallel transport to harmonic maps. In *New Directions in Dirichlet Forms. AMS/IP Stud. Adv. Math.* **8** 49–115. Providence, RI: Amer. Math. Soc. MR1652279

[19] Kendall, W.S. (2001). Symbolic Itô calculus in AXIOM: An ongoing story. *Stat. Comput.* **11** 25–35. MR1837142

[20] Kendall, W.S. (2007). Coupling all the Lévy stochastic areas of multidimensional Brownian motion. *Ann. Probab.* **35** 935–953. MR2319712

[21] Kendall, W.S. (2009). Brownian couplings, convexity, and shy-ness. *Electron. Commun. Probab.* **14** 66–80. MR2481667

[22] Kendall, W.S. (2010). Coupling time distribution asymptotics for some couplings of the Lévy stochastic area. In *Probability and Mathematical Genetics. London Mathematical Society Lecture Note Series* **378** 446–463. Cambridge: Cambridge Univ. Press. MR2744251

[23] Kendall, W.S. and Price, C.J. (2004). Coupling iterated Kolmogorov diffusions. *Electron. J. Probab.* **9** 382–410 (electronic). MR2080604

[24] Kuwada, K. (2009). Characterization of maximal Markovian couplings for diffusion processes. *Electron. J. Probab.* **14** 633–662. MR2486817

[25] Kuwada, K. and Sturm, K.-T. (2007). A counterexample for the optimality of Kendall–Cranston coupling. *Electron. Commun. Probab.* **12** 66–72 (electronic). MR2300216

[26] Lindvall, T. (1982). On coupling of Brownian motions. Technical Report 1982:23, Dept. Mathematics, Chalmers Univ. Technology and Univ. Göteborg.

[27] Lindvall, T. (1991). W. Doeblin 1915–1940. *Ann. Probab.* **19** 929–934. MR1112401

[28] Lindvall, T. (1992). *Lectures on the Coupling Method. Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR1180522

[29] Lindvall, T. and Rogers, L.C.G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872. MR0841588

[30] Pitman, J.W. (1975). One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. in Appl. Probab.* **7** 511–526. MR0375485

[31] Pitman, J.W. (1976). On coupling of Markov chains. *Z. Wahrsch. Verw. Gebiete* **35** 315–322. MR0415775

[32] Revuz, D. and Yor, M. (1991). *Continuous Martingales and Brownian Motion. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Berlin: Springer. MR1083357

[33] Rogers, L.C.G. (1999). Fastest coupling of random walks. *J. London Math. Soc.* (2) **60** 630–640. MR1724813

[34] Rogers, L.C.G. and Williams, D. (1987). *Diffusions*, *Markov Processes*, *and Martingales*. *Vol*. 2. *Itô Calculus*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0921238

[35] Rosenthal, J.S. (1997). Faithful couplings of Markov chains: Now equals forever. *Adv. in Appl. Math*. **18** 372–381. MR1436487

[36] Sigman, K. (2011). Exact simulation of the stationary distribution of the FIFO M/G/c queue. *J. Appl. Probab*. **48A** 209–213. MR2865627

[37] Smith, A. (2014). A Gibbs sampler on the n-simplex. *Ann. Appl. Probab*. **24** 114–130.

[38] Sverchkov, M.Yu. and Smirnov, S.N. (1990). Maximal coupling for processes in $D[0, \infty]$. *Dokl. Akad. Nauk SSSR* **311** 1059–1061. MR1072656

[39] Thorisson, H. (1994). Shift-coupling in continuous time. *Probab. Theory Related Fields* **99** 477–483. MR1288066

[40] Thorisson, H. (2000). *Coupling*, *Stationarity*, *and Regeneration*. *Probability and Its Applications* (*New York*). New York: Springer. MR1741181

[41] Warren, J. and Yor, M. (1998). The Brownian burglar: Conditioning Brownian motion by its local time process. In *Séminaire de Probabilités*, *XXXII*. *Lecture Notes in Math*. **1686** 328–342. Berlin: Springer. MR1655303