Electronic Journal of Probability

Vol. 9 (2004), Paper no. 13, pages 382-410.

Journal URL
http://www.math.washington.edu/~ejpecp/

Coupling Iterated Kolmogorov Diffusions

Wilfrid S. Kendall and **Catherine J. Price**

Department of of Statistics, University of Warwick, UK
w.s.kendall@warwick.ac.uk and cprice@lehman.com

Abstract

The *Kolmogorov (1934) diffusion* is the two-dimensional diffusion generated by real Brownian motion B and its time integral $\int B\,\mathrm{d}\,t$. In this paper we construct successful co-adapted couplings for iterated Kolmogorov diffusions defined by adding iterated time integrals $\int\int B\,\mathrm{d}\,s\,\mathrm{d}\,t, \ldots$ as further components to the original Kolmogorov diffusion. A Laplace-transform argument shows it is not possible successfully to couple all iterated time integrals at once; however we give an explicit construction of a successful co-adapted coupling method for $(B, \int B\,\mathrm{d}\,t, \int\int B\,\mathrm{d}\,s\,\mathrm{d}\,t)$; and a more implicit construction of a successful co-adapted coupling method which works for finite sets of iterated time integrals.

Keywords: BROWNIAN MOTION, CO-ADAPTED COUPLING, EXOTIC COUPLING, KOLMOGOROV DIFFUSION, LOCAL TIME, NON-ADAPTED COUPLING, SHIFT-COUPLING, SKOROKHOD CONSTRUCTION, STOCHASTIC INTEGRAL

AMS 2000 Mathematics Subject Classification: 60J65, 60H10

Submitted to EJP on August 29, 2003. Final version accepted on April 19, 2004.

1 Introduction

The *Kolmogorov (1934) diffusion* is the two-dimensional diffusion generated by real Brownian motion B and its time integral $\int B \, \mathrm{d}\, t$. Analytic studies of distribution and winding rate about $(0, 0)$ have been carried out by McKean (1963). More recent workers (Lachal 1997; Khoshnevisan and Shi 1998; Groeneboom, Jongbloed, and Wellner 1999; Chen and Li 2003) have considered growth asymptotics, distribution under conditioning, and small ball probabilities. Ben Arous et al. (1995) showed that $(B, \int B \, \mathrm{d}\, t)$ can be successfully *coupled co-adaptedly*, meaning that for any two different starting points (a_1, a_2) and (b_1, b_2) it is possible to construct random processes $(A, \int A \, \mathrm{d}\, t)$ and $(B, \int B \, \mathrm{d}\, t)$ begun at (a_1, a_2) and (b_1, b_2) respectively, adapted to the same filtration and such that A and B are real Brownian motions with respect to this filtration, which *couple successfully* in the sense that $A_T = B_T$ and $a_2 + \int_0^T A \, \mathrm{d}\, t = b_2 + \int_0^T B \, \mathrm{d}\, t$ for some random but finite time T. The *iterated Kolmogorov diffusion* is obtained by adding (perhaps a finite number of) further iterated time integrals $\int\int B \, \mathrm{d}\, s \, \mathrm{d}\, t, \ldots$ as components, and the object of this note is to study its coupling properties.

There are many different kinds of coupling: *co-adapted* or *Markovian coupling* as described above, *co-adapted time-changed coupling*, which relaxes the filtration requirements to permit random time-changes (an example is to be found in Kendall 1994); *non-adapted coupling*, which lifts the filtration requirement; and finally *shift-coupling*, which relaxes the coupling requirement to permit coupling up to a random time (Aldous and Thorisson 1993). Elementary martingale arguments show a diffusion cannot be successfully coupled if there exist non-trivial bounded functions which are parabolic (space-time harmonic) with respect to the diffusion; more generally a diffusion cannot be successfully shift-coupled if there exist non-trivial bounded functions which are harmonic. The converse statements are also true: absence of non-constant parabolic functions means there exist successful non-adapted couplings (Griffeath 1975; Goldstein 1979), and absence of non-constant harmonic functions means there exist successful shift-couplings (Aldous and Thorisson 1993).

Co-adapted couplings are generally less powerful than non-adapted couplings, but can provide significant links to mathematical notions such as curvature. For example Kendall (1986) describes a co-adapted coupling construction for Brownian motion on Cartan-Hadamard manifolds of negative curvature bounded above away from zero, and shows that there is no successful co-adapted coupling. If it could be shown in this case that successful non-adapted coupling implied existence of a successful co-adapted coupling, then one could use the link with parabolic functions to deduce that all such manifolds must support non-constant bounded parabolic functions; this question from Riemannian geometry is currently open! Furthermore, it is typically much easier to construct co-adapted couplings when they do exist; a matter of major significance when using coupling to explore convergence of a Markov chain to equilibrium (when using Markov chains as components of approximate counting algorithms as expounded in Jerrum 2003; or when implementing Coupling from the Past as in Propp and Wilson 1996). Burdzy and Kendall (2000) explore the difference between non-adapted and co-adapted couplings; see also Hayes and Vigoda (2003), who describe a non-adapted variation on an adapted coupling which provides better bounds for mixing in a particular graph algorithm.

In this paper we extend the results of Ben Arous et al. (1995) for $(B, \int B \, \mathrm{d}\, t)$, giving an explicit construction for a successful co-adapted coupling at the level of the twice-iterated time integral (Theorem 3.5). We also give an implicit construction for successful co-adapted couplings for higher-order iterated Kolmogorov diffusions (Theorem 5.14), and note that it is impossible successfully to couple all iterated time integrals simultaneously (Theorem 6.1).

Acknowledgements: this work was supported by the EPSRC through an earmarked studentship for CJP. We are grateful to Dr Jon Warren and to Dr Sigurd Assing for helpful discussions.
Part of this work was carried out while the first author was visiting the Institute for Mathematical Sciences, National University of Singapore in 2004. The visit was supported by the Institute.
Finally, we are grateful to a referee, who asked an astute question which led to a substantial strengthening of the results of this paper.

2 Parabolic functions and harmonic functions

We begin by sketching some general mathematical considerations. It is possible to derive information about the existence or otherwise of couplings from analytic considerations, albeit in a rather non-constructive fashion. The existence of a successful non-adapted coupling is known to be equivalent to the nonexistence of non-constant parabolic functions (Griffeath 1975; Goldstein 1979). The same is true if we replace “non-adapted coupling” and “parabolic” by “shift-coupling” and “harmonic” (Aldous and Thorisson 1993). (These papers consider the discrete-time case: the technical issue of moving to continuous time is dealt with for example in Thorisson 2000.)

In general it is known that manifolds which are (for example) unimodular solvable Lie groups will not carry non-constant bounded harmonic functions (Lyons and Sullivan 1984; Kaĭmanovich 1986; Leeb 1993). The iterated Kolmogorov diffusion can be viewed as a Brownian motion on a nilpotent Lie group, so we can deduce the existence of successful *shift-couplings* for the iterated Kolmogorov diffusion.

We are concerned here with successful couplings rather than successful shift-couplings, corresponding to parabolic functions rather than harmonic functions. However Cranston and Wang (2000, §3, Remark 3) show that a parabolic Harnack inequality holds for left-invariant diffusions on unimodular Lie groups (and therefore successful shift-couplings exist for such diffusions if and only if successful non-adapted couplings exist). It suffices to indicate how the iterated Kolmogorov diffusion can be viewed as such a diffusion. We outline the required steps.

First observe that there is a homomorphism of the semigroup of paths B under concatenation into the quotient *group* which identifies paths with the same time-length and the same endpoints and iterated time integrals up to order n. The resulting group is graded by a degree defined inductively by time-integration, and is nilpotent with this grading. It is a Lie group, since it can be coordinatized smoothly by t, B_t, and n iterated integrals of the form $\int_0^t B \,\mathrm{d}\,u$, $\ldots$, $\int_0^t \ldots \int_0^v B \,\mathrm{d}\,w \ldots \mathrm{d}\,u$; evolution of B generates the required left-invariant diffusion. Nilpotent Lie groups are unimodular (Corwin and Greenleaf 1990), so the Cranston and Wang (2000) work applies.

Thus at a rather abstract and indirect level we know it is possible to construct successful non-adapted couplings for the iterated Kolmogorov diffusion. However in the following we will show how to construct successful *co-adapted couplings*; while our general construction (§5) is not completely explicit, nevertheless it is much more direct than the above, as well as possessing the useful co-adapted property.

3 Explicit co-adapted coupling for the twice-iterated Kolmogorov diffusion

We now describe a constructive approach to successful *co-adapted* coupling of Brownian motion and its first two iterated integrals: B, $\int B \,\mathrm{d}\,t$ and $\int\int B \,\mathrm{d}\,s \,\mathrm{d}\,t$. In later sections we will show how to deal with higher-order iterated integrals.

We use the conventional probabilistic language of "event A_n happens eventually in n" to mean, almost surely A_n occurs for all but finitely many n (in measure-theoretic terms this corresponds to the assertion that the event $\bigcup_n \bigcap_{m \geq n} A_n$ has null complement).

3.1 Case of first integral

Coupling of the first two iterated integrals is based on the Ben Arous et al. (1995) coupling construction for $(B, \int B \, \mathrm{d}\, t)$; we begin with a brief description of this in order to establish notation.

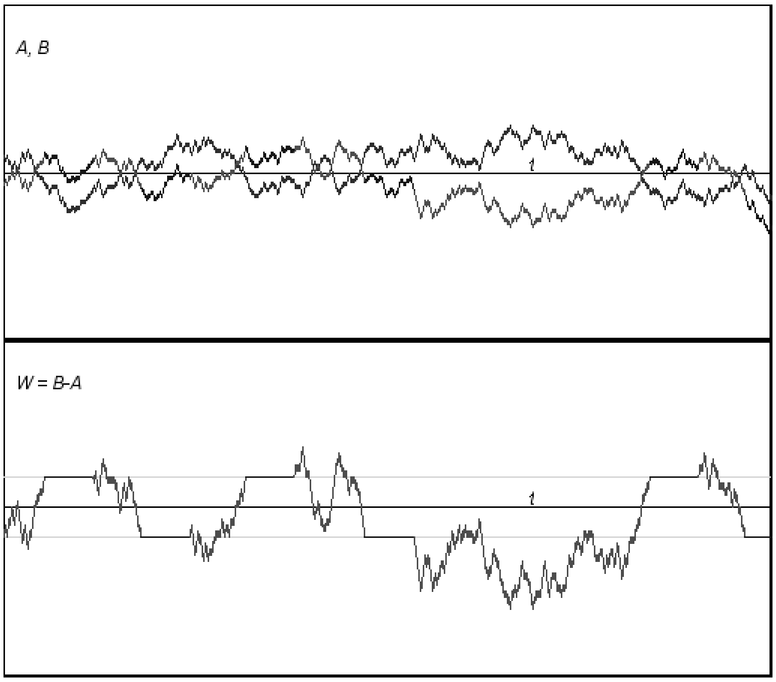

Figure 1: Plots of (a) two coupled Brownian motions A and B, (b) the difference $W = B - A = B(0) - A(0) + \int (J-1) \, \mathrm{d}\, A$. The coupling control J switches between values $+1$ ("synchronous coupling") and -1 ("reflection coupling"). In the figure, switches to fixed periods of $J = +1$ are triggered by successive crossings of ± 1 by W.

Co-adapted couplings are built on two co-adapted Brownian motions A and

B begun at different locations $A(0)$ and $B(0)$: we shall suppose they are related by a stochastic integral $B = B(0) + \int J \, \mathrm{d}\, A$, where J is a piece-wise constant ± 1-valued adapted random function. The coupling is defined by specifying J:

$$W \quad = \quad B - A \quad = \quad B(0) - A(0) + \int (J-1) \, \mathrm{d}\, A \, , \tag{1}$$

so that W is constant on intervals where $J = 1$ (*holding intervals*), and evolves as Brownian motion run at rate 4 on intervals where $J = -1$ (intervals in which W is run at *full rate*). The coupling is illustrated in Figure 1.

So our coupling problem is reduced to a stochastic control problem: how should one choose adapted J so as to control W and $V = V(0) + \int W \, \mathrm{d}\, t$ to hit zero simultaneously?

We start by noting that the trajectory (W, V) breaks up into *half-cycles* according to successive alternate visits to the positive and negative rays of the axis $V = 0$. (We can assume $V(0) = 0$ without loss of generality; we can manipulate W and V to this end using an initial phase of controls!) We adopt a control strategy as follows: if the n^{th} half-cycle begins at $W = \pm a_n$ for $a_n > 0$ then we compute a level b_n depending on a_n, with $b_n \leq a_n \leq \kappa b_n$ for some fixed $\kappa > 1$, and run this half-cycle of W at full rate ($J = -1$) until W hits $\mp b_n$ or the half-cycle ends. If W hits $\mp b_n$ before the end of the half-cycle then we start a holding interval ($J = 1$) until V hits zero, so concluding the half-cycle. Set a_{n+1} to be the absolute value of W at the end of the half-cycle. We will call the holding interval the *Fall* of the half-cycle and will refer to the initial component as the *Brownian component* or *BrC*. The construction is illustrated in Figure 2.

With appropriate choices for the a_n and b_n, it can be shown that this control forces (W, V) almost surely to converge to $(0, 0)$ in finite time. To see this, note the following. By the reflection principle applied to a Brownian motion B begun at 0,

$$\mathbb{P}\left[\textit{BrC} \text{ duration } \geq t_n\right] \quad \leq \quad \mathbb{P}\left[2|B(t_n)| \leq a_n + b_n\right] \quad \leq \quad \frac{a_n + b_n}{\sqrt{2\pi t_n}} \, . \tag{2}$$

Simple dynamical arguments allow us to control the duration of the Fall:

$$\textit{Fall} \text{ duration} \quad \leq \quad \frac{\max V \text{ during } \textit{BrC}}{b_n} \quad \leq \quad \frac{(\textit{BrC} \text{ duration}) \times (\max W \text{ during } \textit{BrC})}{b_n} \, , \tag{3}$$

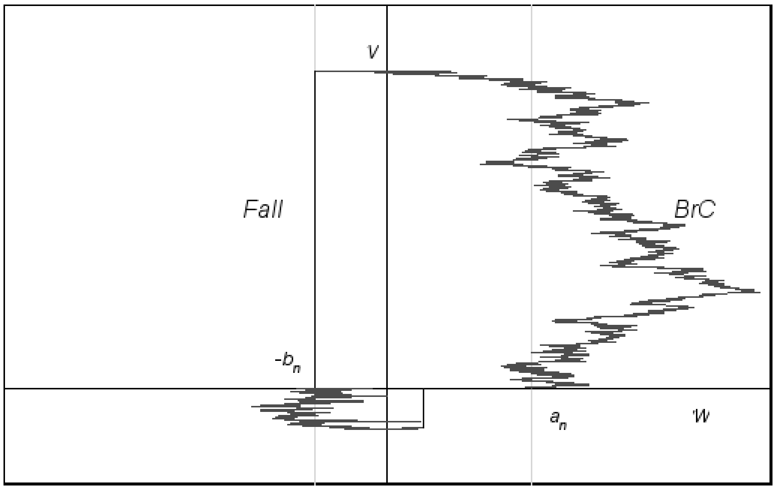

Figure 2: Illustration of two half-cycles for the case $b_n = a_n/2$, $\kappa = 2$, labelling *Fall* and *BrC* for first half-cycle.

which we can combine with the following (for $x_n > 0$):

$$\begin{aligned}\mathbb{P}\left[(\max W \text{ during } \textit{BrC}\,) \geq x_n\right] \quad &= \quad \mathbb{P}\left[2B + a_n \text{ hits } x_n \text{ before } -b_n\right] \\ &= \quad \frac{a_n + b_n}{a_n + b_n + x_n}\,.\end{aligned} \tag{4}$$

We now use a Borel-Cantelli argument to deduce that

$$\text{Duration of half-cycle } n \quad \leq \quad \left(1 + \frac{x_n}{b_n}\right) t_n \tag{5}$$

for all sufficiently large n, so long as

$$\sum_n \frac{a_n}{\sqrt{t_n}} < \infty\,, \qquad\qquad \sum_n \frac{1}{1 + x_n/(2a_n)} < \infty\,. \tag{6}$$

(Bear in mind, we have stipulated that $b_n \leq a_n \leq \kappa b_n$.) Now this convergence is ensured by setting $\sqrt{t_n} = x_n = a_n n^{1+\alpha}$ for some $\alpha > 0$, in which case we obtain

$$\text{Duration of half-cycle } n \quad \leq \quad \left(1 + \frac{\kappa x_n}{a_n}\right) t_n \quad \leq \quad \left(1 + \kappa n^{1+\alpha}\right) a_n^2 n^{2+2\alpha}\,. \tag{7}$$

If we arrange for $a_n \leq \kappa/n^{2+\beta}$ then the sum of this over n converges, since we can choose $\alpha < 2\beta/3$. Thus we have proved the following, which is a trivial generalization of Ben Arous et al. (1995, Theorem 2.1):

Theorem 3.1 *Suppose the evolution of $(W, \int W\, dt)$ is divided into half-cycles as described above: if the n^{th} half-cycle begins at $W = \pm a_n$, then it is run at full rate till W hits $\mp b_n$ and then allowed to fall to the conclusion of the half-cycle. (The fall phase is omitted if the half-cycle concludes before W hits $\mp b_n$.) Our control consists of choosing the b_n; so long as $a_n/\kappa \leq b_n \leq \min\{a_n, 1/n^{2+\beta}\}$ for all sufficiently large n for some constants κ and $\beta > 0$, then $(W, \int W\, dt)$ converges to $(0, 0)$ in finite time.*

Remark 3.2 By definition of a_n we know $a_n \leq b_{n-1} \leq 1/(n-1)^{2+\beta}$, so it is feasible to choose b_n such that $a_n/\kappa \leq b_n \leq \min\{a_n, 1/n^{2+\beta}\}$ for all large n.

Remark 3.3 Note that a_n is determined by the location of W at the end of half-cycle $n-1$.

Remark 3.4 We can assume the initial conditions $W_0 = 1$, $V_0 = 0$ (otherwise we can run the diffusion at full rate till V hits zero, as can be shown to happen almost surely, then re-scale accordingly). It then suffices to set $b_n = \min\{a_n, 1/(n+1)^{2+\beta}\}$. However this is not the only option; for example Ben Arous et al. (1995) use $b_n = a_n/2$. Note, in either case we find $a_{n+1} \leq b_n \leq a_n \leq \kappa b_n$ for $\kappa = 2$.

3.2 Controlling two iterated integrals

Inspection of the above control strategy reveals some flexibility which was not exploited by Ben Arous et al. (1995); in the n^{th} half-cycle there is a time T_n at which W first hits 0, and we may then *hold* $W = 0$ constant (by setting $J = 1$) and so delay for a time C_n, without altering either W or $V = \int W\, \mathrm{d}t$. We may choose C_n as we wish without jeopardizing convergence of (W, V) to $(0, 0)$. This flexibility allows us to consider controlling $U = \int\int W\, \mathrm{d}s\, \mathrm{d}t$ as follows: we hold at T_n for a duration long enough to force U, V to have the same sign:

$$C_n \quad = \quad \max\left\{0, \frac{-\int\int W\, \mathrm{d}s\, \mathrm{d}t}{\int W\, \mathrm{d}t}\right\} \quad = \quad \max\left\{0, -\frac{U}{V}\right\}. \tag{8}$$

The event $[V(T_n) = 0]$ turns out to be of probability zero, since it can only happen if T_n occurs at the very start of the half-cycle, which in turn happens only if W hits zero exactly at the end of the previous half-cycle; that this is a null event will be a weak consequence of the lower bound at Inequality (16) below. The construction is illustrated in Figure 3.

Suppose $a_n \leq \kappa/n^{2+\beta}$ as in the previous subsection. If we can show that $\sum_n C_n < \infty$ then this strategy results in (W, V) tending to $(0, 0)$ in finite time ζ,

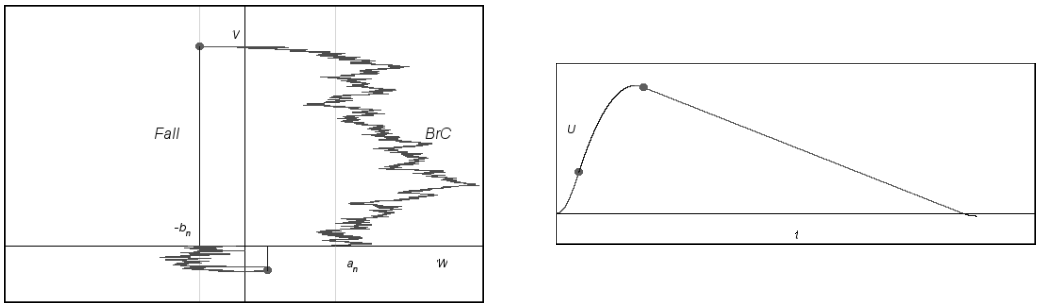

Figure 3: Two consecutive half-cycles for the case $b_n = a_n/2$, together with a graph of U against time. The disks signify time points at which there is an option to *hold* the diffusion to allow U to change sign if required.

with U hitting zero in infinitely many half-cycles accumulating at ζ and therefore also converging to 0 at ζ. To fix notation, let us suppose W is positive at the start of the half-cycle in question. This ensures $V > 0$ at time T_n. So the issue at hand is to control C_n by determining what makes $-U = -\int\int W \,\mathrm{d}\, s \,\mathrm{d}\, t$ large, and what makes $V = \int W \,\mathrm{d}\, t$ small at time T_n.

Consider $-U$ at time T_n. At the start of the half-cycle we know $V = 0$ and $W = a_n$, so subsequent contributions make $-U$ more negative and need not detain us. At the start of the *previous* half-cycle U will be non-negative. Consequently an upper bound for $-U$ at time T_n is given by

$$\frac{(\text{Duration of half-cycle } n-1)^2}{2} \times (-\text{minimum value of } W \text{ over half-cycle } n-1)\,,$$

thus (given the work of §3.1) we may suppose that, eventually in n, at time T_n the quantity $-U$ is bounded above by

$$\frac{1}{2}\left(\left(1+\frac{x_{n-1}}{b_{n-1}}\right) t_{n-1}\right)^2 \times x_{n-1} \quad = \quad \frac{1}{2}\left(1+\kappa(n-1)^{1+\alpha}\right)^2 a_{n-1}^5 (n-1)^{5+5\alpha}\,. \tag{9}$$

Now apply a Borel-Cantelli argument and the reflection principle to show that eventually in n the Brownian component takes time at least $a_n^2/(4n^{2+2\alpha})$ in travelling from a_n to $a_n/2$. We deduce almost surely for all sufficiently large n at time T_n it must be the case that $V = \int W \,\mathrm{d}\, t$ exceeds

$$(a_n/2) \times \text{time to move from } a_n \text{ to } a_n/2 \quad \geq \quad \frac{a_n^3}{8n^{2+2\alpha}}\,, \tag{10}$$

Thus C_n is bounded above, eventually in n, by

$$4\left(1+\kappa(n-1)^{1+\alpha}\right)^2(n-1)^{5+5\alpha}n^{2+2\alpha}\times\left(\frac{a_{n-1}}{a_n}\right)^3 a_{n-1}^2\,. \tag{11}$$

This leads to the crux of the argument; we need an eventual upper bound on the ratio a_{n-1}/a_n.

First note that the lower bound of Inequality (10), applied to the $(n-1)^{\text{st}}$ cycle, shows that eventually in n

$$-\int^{T_{n-1}} W\,\mathrm{d}\,t \quad\geq\quad \frac{a_{n-1}^3}{8(n-1)^{2+2\alpha}}\,. \tag{12}$$

So it suffices to obtain a suitable lower bound on a_n, the value of W at the end of half-cycle $n-1$, in terms of $-\int^{T_{n-1}} W\,\mathrm{d}\,t$ and holding eventually in n. Moreover we may ignore the *Fall* component of half-cycle $n-1$ so long as the lower bound is smaller than b_{n-1}, and treat W over the whole of this half-cycle as a Brownian motion of rate 4.

We now introduce a discontinuous time-change based on the continuous (but non-monotonic) additive functional $V(t)=\int W\,\mathrm{d}\,s$: condition on $\int^{T_{n-1}} W\,\mathrm{d}\,t = -u_{n-1}$ and set

$$\sigma(u) \quad=\quad \inf\left\{s\geq 0: V(T_n+s)-V(T_n)>u-u_{n-1}\right\}$$

for $0\leq u\leq u_{n-1}$. We set $\widetilde{Z}(u)=W(\sigma(u)$ and use standard time-change arguments to show $\mathrm{d}\,u = W(\sigma(u))\,\mathrm{d}\,\sigma$ so $\mathrm{d}\,\sigma/\,\mathrm{d}\,u = 1/W(\sigma(u)) = 1/\widetilde{Z}(u)$. Consequently, on time intervals throughout which $\widetilde{Z}>0$,

$$\mathrm{d}\,\widetilde{Z}(u) \quad=\quad \frac{2}{\sqrt{\widetilde{Z}}}\,\mathrm{d}\,\widetilde{B} \quad \text{when } \widetilde{Z}>0 \tag{13}$$

for a new standard Brownian motion $\widetilde{B}$. The time-changed process $\widetilde{Z}$ is illustrated in Figure 4. Note that $\widetilde{Z}$ must be non-negative.

A nonlinear transformation of scale $Z=\widetilde{Z}^{3/2}/3$ produces a Bessel($\frac{4}{3}$) process Z in time intervals throughout which $\widetilde{Z}$ (equivalently Z) is positive: by Itô's formula the stochastic differential equation

$$\begin{aligned}\mathrm{d}\,Z \quad=\quad \mathrm{d}\left(\frac{1}{3}\widetilde{Z}^{3/2}\right) \quad&=\quad \frac{1}{2}\widetilde{Z}^{1/2}\,\mathrm{d}\,\widetilde{Z}+\frac{1}{8}\frac{1}{\sqrt{\widetilde{Z}}}(\mathrm{d}\,\widetilde{Z})^2\\ &=\quad \mathrm{d}\,\widetilde{B}+\frac{1}{6}\frac{1}{Z}\,\mathrm{d}\,u\,,\end{aligned}$$

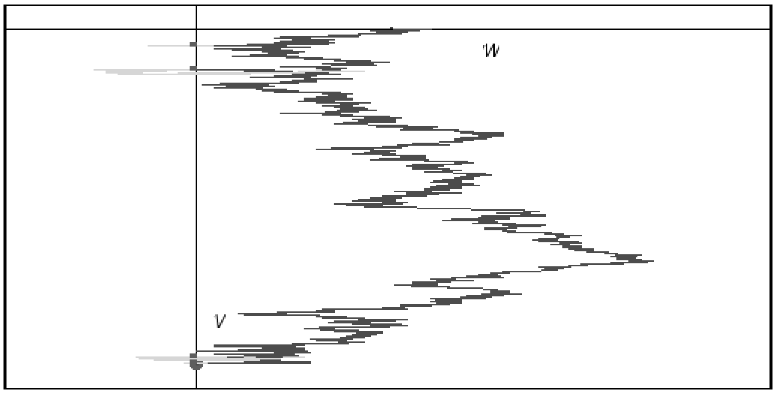

Figure 4: Discontinuous time-change for W based on $V = \int W \, \mathrm{d}\, t$. The effect of the time-change is to delete the loops extending into $W > 0$, and to continue deletion till $V = \int W \, \mathrm{d}\, t$ re-attains its minimum, thus generating discontinuities (some of which are indicated in the figure by small dots on the V-axis) in the time-changed process $\widetilde{Z}$, which follows the red /dark trajectory.

holds in intervals for which $Z > 0$.

Now observe that the zero-set $\{u : Z(u) = 0\}$ is almost surely a null-set. For certainly the Brownian zero-set $\{t : W(t) = 0\}$ is almost surely null, and $\{V(t) : W(t) = 0\}$ is then almost surely null by Sard's lemma, since V is a C^1 function with derivative W (indeed this is an easy exercise in this simple context). But $\{u : Z(u) = 0\} = \{u : \widetilde{Z}(u) = 0\}$ is a subset of $\{V(t) : W(t) = 0\}$.

Because $\{u : Z(u) = 0\}$ is almost surely a null-set, it follows that the stochastic integral

$$B(u) \quad = \quad \int_0^u \mathbb{I}\left[Z > 0\right] d\widetilde{B} \tag{14}$$

(using Equation (13) to construct $\widetilde{B}$ when $Z > 0$) defines a Brownian motion. Furthermore we can apply limiting arguments to write

$$Z \quad = \quad B + \frac{1}{6} \int \mathbb{I}\left[Z > 0\right] \frac{1}{Z} \, \mathrm{d}\, u + H \tag{15}$$

where H is the non-decreasing pure jump process

$$H(u) \quad = \quad \sum_{v \leq u} \frac{1}{3} \left(W(\sigma(u)) - W(\sigma(u-))\right)^{3/2} \, ,$$

which is constant away from $Z = 0$.

Using the theory of the Skorokhod construction (El Karoui and Chaleyat-Maurel 1978), we now derive the comparison

$$Z \quad \geq \quad B + L$$

where L is local time of B at 0, so that $B + L$ is reflected Brownian motion. (H is discontinuous, but the argument of the Skorokhod construction applies so long as H is non-decreasing.) Consequently

$$\mathbb{P}\left[\frac{1}{3}\widetilde{Z}^{3/2}(u_{n-1}) \geq x\right] \quad \leq \quad \mathbb{P}\left[B(u_{n-1}) \geq x\right]$$

and so we can deduce from a Borel-Cantelli argument and the lower bound of Inequality (12) that the following holds eventually in n:

$$a_n \quad \geq \quad \left(\frac{|\int^{T_{n-1}} W \,\mathrm{d}\, t|}{(n-1)^{2+2\alpha}}\right)^{1/3} \quad \geq \quad a_{n-1}\left(\frac{1}{8(n-1)^{4+4\alpha}}\right)^{1/3} \tag{16}$$

for $\alpha > 0$. Thus eventually in n

$$C_n \quad \leq \quad 32\left(1 + (n-1)^{1+\alpha}\right)^2 (n-1)^{9+9\alpha} n^{2+2\alpha} \times a_{n-1}^2$$

and so we can arrange for $\sum_n C_n < \infty$ so long as we choose $a_n \leq 1/n^{7+\beta}$ for $\beta > 0$.

We state this as a theorem:

Theorem 3.5 *The modified strategy described at the head of this subsection (hold at each T_n till $\int\int W\, d\, s\, d\, t$ is zero or has the same sign as $\int W\, d\, t$) produces convergence of*

$$(W, \int W\, d\, t, \int\int W\, d\, s\, d\, t)$$

to $(0, 0, 0)$ in finite time so long as the conditions of Theorem 3.1 are augmented by the following: half-cycle n begins at $W = \pm a_n$ for $a_n \leq 1/n^{7+\beta}$ for $\beta > 0$ and for all sufficiently large n. (This can be arranged by choosing b_n in the range $a_n/\kappa \leq b_n \leq \min\{a_n, 1/n^{7+\beta}\}$ for constants κ and $\beta > 0$.)

Remark 3.6 The choice $b_n = a_n/2$ of Ben Arous et al. (1995) will suffice.

Remark 3.7 It is possible to obtain a modest gain on the above by recognizing that the hitting time of Brownian motion on -1 has the same distribution as the inverse square of a standard normal random variable. This argument permits replacement of Equation (10) by

$$\int^{T_n} W\,\mathrm{d}\,t \quad \geq \quad \text{constant} \times a_n^3 \frac{1}{\sqrt{n}}\,.$$

Remark 3.8 The elementary comparison approach above can of course be replaced by arguments employing the exact computations of McKean (1963).

Remark 3.9 The method described here (control coupling of higher-order iterated integrals by judicious waits at $W = 0$) appears to deliver effective control of just one higher-order iterated integral in addition to W, $V = \int W\,\mathrm{d}\,t$. Attempts to control more than one higher-order iterated integral seem to lead to problems of propagation of over-correction from one half-cycle to the next. We therefore turn to a rather different, less explicit, approach in the remainder of the paper.

4 Reduction to non-iterated time integrals

Before considering the problem of coupling more than two iterated time integrals, we first reformulate the coupling problem in terms of integrals of the form $\int \frac{t^m}{m!} B(t)\,\mathrm{d}\,t$ rather than the less amenable iterated time integrals of above. We begin with some notation. Suppose W is defined as the difference between two co-adapted coupled Brownian motions, as in §3.1. Then we set $W = W^{(0)} = B - A$ and define the first N iterated time integrals inductively by

$$\begin{aligned}
W^{(1)} &\quad = \quad W^{(1)}(0) + \int W^{(0)}\,\mathrm{d}\,t\,, \\
W^{(2)} &\quad = \quad W^{(2)}(0) + \int W^{(1)}\,\mathrm{d}\,t\,, \\
&\quad \ldots \\
W^{(N)} &\quad = \quad W^{(N)}(0) + \int W^{(N)}\,\mathrm{d}\,t\,.
\end{aligned} \tag{17}$$

(Note that we have allowed for arbitrary initial conditions $W^{(1)}(0)$ *et cetera*.)

If $W^{(0)}(0) = W^{(1)}(0) = W^{(2)}(0) = \ldots = W^{(N)}(0) = 0$, so the initial conditions all vanish, then we find by exchange of integrals that

$$
\begin{aligned}
W^{(n)}(T) &= \int_0^T W^{(n-1)}(t)\,\mathrm{d}\,t \quad = \quad \int_0^T \left(\int_0^t W^{(n-2)}(s)\,\mathrm{d}\,s\right)\mathrm{d}\,t \\
&= \int_0^T \left(\int_s^T W^{(n-2)}(s)\,\mathrm{d}\,t\right)\mathrm{d}\,s \quad = \quad \int_0^T (T-s)W^{(n-2)}(s)\,\mathrm{d}\,s \\
&= \quad \ldots \quad = \quad \int_0^T \frac{(T-s)^{n-1}}{(n-1)!} W^{(0)}(s)\,\mathrm{d}\,s\,.
\end{aligned}
$$

Binomial expansion leads to the following:

Lemma 4.1 *Suppose* $W^{(0)}(0) = W^{(1)}(0) = W^{(2)}(0) = \ldots = W^{(N)}(0) = 0$*, and* J *is a given adapted control, and* ζ *is a given stopping time. Then*

$$
W^{(0)}(\zeta) = 0, \qquad W^{(1)}(\zeta) = 0, \qquad W^{(2)}(\zeta) = 0, \quad \ldots, \quad W^{(N)}(\zeta) = 0
$$

if and only if

$$
\begin{aligned}
W(\zeta) \quad &= \quad 0\,, \quad \int_0^\zeta W(t)\,dt \;=\; 0\,, \quad \int_0^\zeta tW(t)\,dt \;=\; 0\,, \\
&\ldots\,, \quad \int_0^\zeta \frac{t^{N-1}}{(N-1)!} W(t)\,dt \;=\; 0\,.
\end{aligned}
$$

If the iterated time integrals have non-zero initial values then we can reduce to the case of zero initial values by supposing W is deterministically extended backwards in time to time -1, with corresponding generalization of Equation (17). By a simple argument using orthogonal Legendre polynomials P_n on $[-1,1]$, we can choose $W|_{[-1,0]}$ to produce $W^{(0)}(-1) = W^{(1)}(-1) = W^{(2)}(-1) = \ldots = W^{(N)}(-1) = 0$, and $W^{(0)}(0) = a_0$, $W^{(1)}(0) = a_1$, $W^{(2)}(0) = a_2$, …, $W^{(N)}(0) = a_N$ as required. For if

$$
W(t) \quad = \quad \sum_{n=0}^{N+1} b_n P_n(2t+1)
$$

for $-1 \le t \le 0$ then

$$
\int_{-1}^0 P_n(2t+1)W(t)\,\mathrm{d}\,t \quad = \quad \frac{1}{2} b_n \int_{-1}^1 P_n(t)^2\,\mathrm{d}\,t\,.
$$

Expanding the Legendre polynomials and adapting the argument leading to Lemma 4.1, we can find $a_1, \ldots, a_N$ in terms of $b_0, b_1, \ldots, b_{N-1}$ by solving a triangular linear system of equations. Finally, b_N and b_{N+1} may be fixed by the requirement that $W(0) = a_0 = \sum_{n=0}^{N+1} b_n P_n(1)$ and $0 = W(-1) = \sum_{n=0}^{N+1} b_n P_n(0)$ (note that Legendre polynomials do not vanish at their end-points, and are odd or even functions according to whether their order is odd or even!).

This allows us to use Lemma 4.1 to deduce the required reduction:

Lemma 4.2 *It is possible to use adapted controls J to ensure*

$$W^{(0)}(\zeta) = 0, \qquad W^{(1)}(\zeta) = 0, \qquad W^{(2)}(\zeta) = 0, \quad \ldots, \quad W^{(N)}(\zeta) = 0$$

at some stopping time ζ (depending on initial values $W^{(n)}(0)$) if and only if it is also possible to use adapted controls to ensure

$$\begin{aligned} W(\zeta) \quad &= \quad 0\,, \\ b_0 + \int_0^\zeta W(t)\,dt \;&=\; 0\,, \quad b_1 + \int_0^\zeta tW(t)\,dt \;=\; 0\,, \\ \ldots\,, \quad b_{N-1} + \int_0^\zeta &\frac{t^{N-1}}{(N-1)!} W(t)\,dt \;=\; 0\,. \end{aligned}$$

(Here the constants b_n depend on the initial conditions $W^{(n)}(0)$ for the iterated integrals).

For we may extend back over $[-1, 0]$, condition on achieving the desired values $\{W^{(n)}(0) : n = 0, \ldots, N\}$, and work with controls J which act only for positive time.

5 Coupling finitely many iterated integrals

To motivate the control strategy required to couple more than two iterated integrals, we consider a discrete analogue to our problem which is in fact a limiting case.

Suppose we choose only to switch between $J = \pm 1$ at instants when W switches over between two constant levels (as illustrated in Figure 1). To aid explanation we temporarily entertain the fiction that the Brownian motions conspire to produce instantaneous switching as soon as J is switched to -1. Then it is a matter of simple integration to compute the effect on integrals of the form

$\int \frac{t^m}{m!}\,\mathrm{d}W$: if we hold $J = +1$ at successive levels ± 1, beginning at $+1$, making switches from ± 1 to ∓ 1 at times $0 = T_0 < T_1 < \ldots < T_r$, then (under the instantaneous switching approximation)

$$\int_0^{T_r} \frac{t^m}{m!}\,\mathrm{d}W \quad = \quad \sum_{k=0}^{r-1} (-1)^k \frac{T_{k+1}^{m+1} - T_k^{m+1}}{(m+1)!}\,. \tag{18}$$

In §5.1 below we show that particular patterns of switching times produces zero effect on integrals up to a fixed order, at least for the discrete analogue. This permits us to eliminate a whole finite sequence of the integrals. (Of course in practice, because switching is not instantaneous, the use of such patterns creates further contributions to the integrals which then must be dealt with in turn!)

It is algebraically convenient to formulate the required patterns using a sequence $S_0, S_1, \ldots S_{2^N-1}$ of values of ± 1 defined recursively in a manner reminiscent of the theory of experimental design. We set

$$\begin{aligned} S_0 \quad &= \quad +1\,, \\ (S_{2^n}, S_{2^n+1}, \ldots, S_{2^{n+1}-1}) \quad &= \quad -(S_0, S_1, \ldots, S_{2^n-1})\,. \end{aligned} \tag{19}$$

Here is the pattern formed by the first sixteen S_n values:

$$+\;-\;-\;+\quad -\;+\;+\;-\quad -\;+\;+\;-\quad +\;-\;-\;+$$

We will be considering perturbations and re-scalings of the deterministic control which applies control $J = -1$ throughout the time interval $[m, m+1)$ till a switch has occurred to level S_m, and then applies $J = 1$ for the remainder of the time interval. The discrete analogue can be viewed as a limiting case under homogeneous (not Brownian!) scaling of space and time. See Figure 5 for an illustration.

5.1 Algebraic properties of the sign sequence

We now prove some simple properties of the sign sequence $S_0, S_1, \ldots S_{2^N-1}$.

Lemma 5.1 *If $b(r)$ is the number of positive bits in the binary expansion of r then*

$$S_r \quad = \quad (-1)^{b(r)}\,.$$

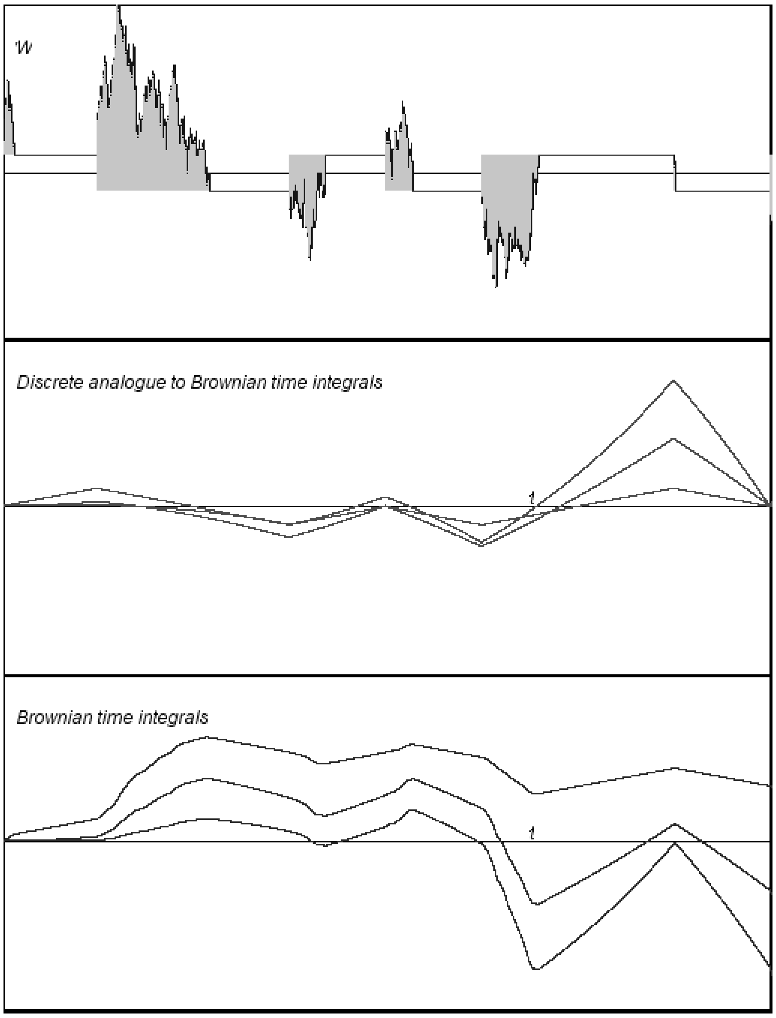

Figure 5: Illustration of the deterministic switching control corresponding to the pattern $+ \ - \ - \ + \ - \ + \ + \ -$. The top panel shows the path of W together with shaded regions indicating discrepancies with the discretized analogue. The middle panel shows the paths of the first three time integrals of the discretized analogue (all of which are brought to 0 at the end of the sequence). The lower panel shows the paths of the first three time integrals of W.

Proof:
Since $b(0) = 0$ this holds for $S_0 = 1$. The recursive definition (19) shows that if the lemma holds for the first 2^n entries in the sequence of S_n then it will also hold for the next 2^n entries. The result follows by induction. $\square$

The proofs of the next three corollaries are immediate from the recursive definition of the S_m's.

Corollary 5.2

$$S_{2^\ell j+a} \quad = \quad S_j S_a \qquad \textit{when } a = 0, 1, \ldots, 2^\ell - 1.$$

Corollary 5.3

$$S_{2m+1} \quad = \quad -S_{2m}$$

Notice the analogue of Corollary 5.3 does *not* hold between S_{2m+2} and S_{2m+1}!

Corollary 5.4

$$\begin{aligned} \sum_{m=0}^{2^n-1} S_m &= 0 && \textit{if } n \geq 1\,, \\ &= 1 && \textit{if } n = 0\,. \end{aligned}$$

These results imply the vanishing of certain sums of low-order powers:

Lemma 5.5

$$\begin{aligned} \sum_{m=0}^{2^N-1} \frac{(m+1)^k}{k!} S_m &= 0 && \textit{if } k < N\,, \\ &= (-1)^N 2^{N(N-1)/2} && \textit{if } k = N\,. \end{aligned} \tag{20}$$

Proof:
Use induction on the level k. If $k = 0$ then Equation (20) follows from the expression for $\sum_{m=0}^{2^N-1} S_m$ in Corollary 5.4. Hence Equation (20) holds at level $k = 0$.

Suppose Equation (20) holds for all levels below level k and suppose $k < N$. Using the recursive construction (19),

$$\begin{aligned}\sum_{m=0}^{2^N-1}\frac{(m+1)^k}{k!}S_m &= \sum_{m=0}^{2^{N-1}-1}\frac{(m+1)^k-(m+1+2^{N-1})^k}{k!}S_m\\ &= -\sum_{u=1}^{k}\frac{(2^{N-1})^u}{u!}\sum_{m=0}^{2^{N-1}-1}\frac{(m+1)^{k-u}}{(k-u)!}S_m\end{aligned}$$

using the binomial expansion and cancelling the $(m+1)^k$ terms. Now we can apply the inductive hypothesis to dispose of terms involving $(m+1)^{k-u}$ for $u > 1$:

$$\sum_{m=0}^{2^N-1}\frac{(m+1)^k}{k!}S_m = -2^{N-1}\sum_{m=0}^{2^{N-1}-1}\frac{(m+1)^{k-1}}{(k-1)!}S_m\,.$$

Thus Equation (20) follows by applying the inductive hypothesis to the right-hand side. □

Remark 5.6 An alternative approach uses generating functions, applying the recursive construction of Equation (19) to show

$$\sum_{k=0}^{\infty}u^k\left(\sum_{m=0}^{2^N-1}\frac{(m+1)^k}{k!}S_m\right) = \left(1-e^{2^{N-1}u}\right)\ldots\left(1-e^{2u}\right)\left(1-e^{u}\right)e^{u}\,.$$

We can now compute the discrete analogue of $\int_0^{2^N}\frac{t^m}{m!}W(t)\,\mathrm{d}\,t$ under the special control which arranges instantaneous switching to level S_m at time m.

Theorem 5.7 *For all* $k < N$

$$\sum_{m=0}^{2^N-1}\frac{(m+1)^{k+1}-m^{k+1}}{(k+1)!}S_m = 0\,. \tag{21}$$

Proof:
Equation (21) follows from Equation (20) by binomial expansion: for example we use $m = (m+1)-1$ to deduce

$$\sum_{m=0}^{2^N-1}\frac{(m+1)^N-m^N}{N!}S_m = \sum_{u=0}^{N-1}\frac{(-1)^{N-u}}{(N-u)!}\sum_{m=0}^{2^N-1}\frac{(m+1)^u}{u!}S_m = 0\,.$$

$\square$

Note that Equation (21) for $S_0, S_1, \ldots, S_{2^N-1}$ is equivalent to Equation (18) with appropriate definitions of the switching times $T_0, T_1, \ldots, T_r$. We now introduce notation for these deterministic times, as they will be basic to our coupling construction.

Definition 5.8 *Define switching times $\widetilde{T}_1, \ldots, \widetilde{T}_{r(N)}$ to be the times m at which $S_{m-1} = \pm 1$ switches to $S_m = \mp 1$, and set $\widetilde{T}_0 = 0$.*

Corollary 5.9 *By induction on $N \geq 1$, since $r(N+1) \geq 2r(N)$,*

$$N \quad \leq \quad r(N) \quad \leq \quad 2^N - 1\,.$$

5.2 Application to the coupling problem

We can now summarize our control strategy for successful coupling. By global analysis (specifically, the inverse function theorem), as long as initial conditions for integrals of order up to $\int \frac{t^{N-1}}{(N-1)!} B \, \mathrm{d}\, t$ are sufficiently small we can obtain a perturbation $T_0^0 = 2^N < T_1^0 < \ldots < T_{r(N)}^0$ of the switching times $\widetilde{T}_1, \ldots, \widetilde{T}_{r(N)}$ (Definition 5.8) which will dispose of these initial conditions by time 2^N. (As will become apparent, scaling arguments can be deployed to deal with larger initial conditions.)

Since the coupled Brownian motions cannot actually produce *instantaneous* transitions between ± 1, the switching activity will have introduced further non-zero contributions to the integrals by time 2^N. So long as these contributions are in turn sufficiently small, we can dispose of them in turn by administering a new control based on switching times $T_0^1 = 2^N < T_1^1 < \ldots < T_{r(N)}^1$ which form a small perturbation of the *scaled* control obtained from $\widetilde{T}_1, \ldots, \widetilde{T}_{r(N)}$ by re-scaling both space and time homogeneously by a factor $1/2$ (*not* by Brownian scaling!) and shifting forwards in time by 2^N. Thus switching now occurs between levels $\pm 1/2$. A key reason for the success of the coupling is that in terms of Brownian scaling there is now twice as much effective time in which to carry out each switch! This means that the probability of all switches completing within their assigned times will increase rapidly to 1.

We can now continue this procedure, disposing of further non-zero contributions by appending further controls using perturbations based on smaller delays and smaller levels. A Borel-Cantelli argument shows there is a positive lower bound on the probability of this infinite sequence completing before a finite time: and moreover the size of the integrals decreasing to zero.

If this fails (because at some stage no small enough perturbation is available, or because a switch fails to complete before its successor is due) then we simply restart the procedure, re-scaling time to ensure existence of the perturbation required initially. Continuing in this manner allows us to deduce that almost surely coupling is eventually successful.

All depends on analyzing the behaviour of perturbations of deterministic controls of the form of Definition 5.8. Consider the map whose coordinates correspond to analogues of time integrals of order less than N:

$$\underline{F}(t_0;\underline{t},\underline{u}) = \underline{F}(t_0;t_1,\ldots,t_N,u_{N+1},\ldots,u_{r(N)}) = (F_0,\ldots,F_{N-1}) \quad (22)$$

$$F_m(t_0;\underline{t},\underline{u}) = \sum_{k=0}^{N-1}(-1)^k\frac{t_{k+1}^{m+1}-t_k^{m+1}}{(m+1)!}+(-1)^N\frac{u_{N+1}^{m+1}-t_N^{m+1}}{(m+1)!}+ \quad (23)$$
$$+\sum_{k=N+1}^{r(N)-1}(-1)^k\frac{u_{k+1}^{m+1}-u_k^{m+1}}{(m+1)!}\,,$$

so $F_m(t_0;\underline{t},\underline{u})$ describes the contribution to $\int \frac{t^m}{m!}W\,\mathrm{d}\,t$ made by instantaneous switching between levels ±1 happening at times t_1, ..., t_N, $u_{N+1}, \ldots u_{r(N)}$, starting at level 1 at time t_0. (Recalling Corollary 5.9, $r(N)\geq N$ if $N\geq 1$.) We can re-write $F_m(t_0;\underline{t},\underline{u})$ as

$$F_m(t_0;\underline{t},\underline{u}) = 2\sum_{k=1}^{N}(-1)^{k-1}\frac{t_k^{m+1}}{(m+1)!}+(\text{further terms involving } t_0,\, u_{N+1},\ldots)\,.$$

We compute the Jacobian for $F(t_0;\underline{t},\underline{u})$ with respect to the arguments t_1, ..., t_N:

$$\det\begin{pmatrix} 2 & 2t_1 & 2\frac{t_1^2}{2!} & \ldots & 2\frac{t_1^{N-1}}{(N-1)!}\\ -2 & -2t_2 & -2\frac{t_2^2}{2!} & \ldots & -2\frac{t_2^{N-1}}{(N-1)!}\\ & & & \ldots & \\ 2(-1)^{N-2} & 2(-1)^{N-2}t_{N-1} & 2(-1)^{N-2}\frac{t_{N-1}^2}{2!} & \ldots & 2(-1)^{N-2}\frac{t_{N-1}^{N-1}}{(N-1)!}\\ 2(-1)^{N-1} & 2(-1)^{N-1}t_N & 2(-1)^{N-1}\frac{t_N^2}{2!} & \ldots & 2(-1)^{N-1}\frac{t_N^{N-1}}{(N-1)!}\end{pmatrix}.$$

This is proportional to a Vandermonde determinant and in fact evaluates to

$$\frac{2^N(-1)^N}{1\cdot 1\cdot 2!\cdot\ldots\cdot(N-1)!}\prod_{1\leq i<j\leq N}(t_j-t_i)\,, \quad (24)$$

which is non-zero so long as the t_i's are distinct. This and the inverse function theorem allows us to assert the following fact:

Lemma 5.10 *The polynomial (hence smooth) map*

$$\underline{t} \;=\; (t_1, \ldots, t_N) \quad \mapsto \quad F(0; \underline{t}, \widetilde{T}_{N+1}, \ldots, \widetilde{T}_{r(N)})$$

is invertible in a neighbourhood of the initial sequence of switching times $\widetilde{T}_0$, $\widetilde{T}_1$, …, $\widetilde{T}_N$ corresponding to switching between levels S_m, … at times m, …. In particular there is $\kappa > 0$ and $\varepsilon' > 0$ such that for all $\varepsilon < \varepsilon'$, if $|W^{(m+1)}(0)| < \varepsilon$ for $m = 0, 1, \ldots, N-1$, then there is a $\kappa\varepsilon$-perturbation $(t_1, \ldots, t_N)$ of $(\widetilde{T}_1, \ldots \widetilde{T}_N)$ with $0 < t_1 < \ldots < t_N < \widetilde{T}_{N+1}$ (hence generating a valid switching strategy) which is such that

$$F_m(0; t_1, \ldots, t_N, \widetilde{T}_{N+1}, \ldots, \widetilde{T}_{r(N)}) \quad = \quad -W^{(m+1)}(0) \qquad \text{for } m = 0, 1, \ldots, N-1.$$

Note further that from Equation (22) and the binomial theorem we have a translation symmetry:

Lemma 5.11

$$F_m(t_0 + s; \underline{t} + s, \underline{u} + s) \quad = \quad \sum_{u=0}^{m} \frac{s^{m-u}}{(m-u)!} F_u(t_0; \underline{t}, \underline{u})$$

while Equation (22) directly yields a scaling property:

Lemma 5.12

$$F_m(\lambda t_0; \lambda \underline{t}, \lambda \underline{u}) \quad = \quad \lambda^{m+1} F_m(t_0; \underline{t}, \underline{u})\,.$$

We need just one more lemma, concerning the behaviour of Brownian motion, before we can state and prove the main coupling result for this section.

Lemma 5.13 *Consider a Brownian motion B started at 0 and run till it hits level $-3/2$ at time S. For fixed constants K_1, K_2*

$$\sum_n \mathbb{P}\left[S > K_1(1+n)^{2+2\alpha} \text{ or } \sup_{[0,S]} B > K_2(1+n)^{2+2\alpha} \right] \quad < \quad \infty\,.$$

Proof:
This follows easily from the reflection principle and elementary Gaussian integral estimates:

$$\begin{aligned}
& \mathbb{P}\left[S > (1+n)^{2+2\alpha} \text{ or } \sup_{[0,S]} B > (1+n)^{2+2\alpha}\right] \\
\leq \quad & 1 - 2\,\mathbb{P}\left[B(1+n)^{2+2\alpha} < -\frac{3}{2}\right] + 2\,\mathbb{P}\left[B((1+n)^{2+2\alpha}) > (1+n)^{2+2\alpha}\right] \\
\leq \quad & \frac{2}{\sqrt{2\pi}} \int_{-\frac{3}{2(1+n)^{1+\alpha})}}^{\frac{3}{2(1+n)^{1+\alpha})}} e^{-u^2/2}\,\mathrm{d}\,u + \frac{2}{\sqrt{2\pi}} \int_{(1+n)^{1+\alpha}}^{\infty} e^{-u^2/2}\,\mathrm{d}\,u \\
\leq \quad & \frac{5}{\sqrt{2\pi}} \frac{1}{(1+n)^{1+\alpha}}\,.
\end{aligned}$$

□

Theorem 5.14 *There is a successful co-adapted coupling for Brownian motion and its first N iterated time integrals.*

Proof:
By the work of §4 this is reduced this to the problem of finding an adapted control $J = \pm 1$ which delivers W such that at a particular stopping time ζ we have $W(\zeta) = W^{(0)}(\zeta) = 0$ and $W^{(m+1)}(0) + \int_0^\zeta \frac{t^m}{m!} W(t)\,\mathrm{d}\,t = 0$ for $m = 0, 1, \ldots, N-1$. Without loss of generality we assume $W(0) = 2$ (for otherwise we can run the control $J = -1$ till this occurs!).

Using Lemma 5.10 and Lemma 5.12, for fixed $\varepsilon > 0$ we can choose C large enough to solve for $(t_1^0, \ldots, t_N^0)$ with $|t_k^0 - \widetilde{T}_k| < \varepsilon$ in

$$F_m(0; Ct_1^0, \ldots, Ct_N^0, C\widetilde{T}_{N+1}, \ldots, C\widetilde{T}_{r(N)}) \quad = \quad -W^{(m+1)}(0) \tag{25}$$

for $m = 0, \ldots, N-1$: carry out this switching strategy over the time period $[0, 2^N C)$ to eliminate the initial conditions.

We now apply the following algorithm iteratively starting at step $k = 1$, and continuing, to reduce the further contributions to the integrals made during previous switching strategies.

Step k: At time $T_0^k = (2 - 2^{1-k})2^N C$, use Lemmas 5.10, 5.11, and 5.12 to determine the solution $(t_1^k, \ldots, t_N^k)$ of

$$F_m(T_0^k; T_0^k + 2^{-k}Ct_1^k, \ldots, T_0^k + 2^{-k}Ct_N^k, T_0^k + 2^{-k}C\widetilde{T}_{N+1}, \ldots, T_0^k + 2^{-k}C\widetilde{T}_{r(N)})$$
$$= -\int_0^{T_0^k} \frac{t^m}{m!} W(t)\,\mathrm{d}\,t$$
$$= -\int_{T_0^{k-1}}^{T_0^k} \frac{t^m}{m!} W(t)\,\mathrm{d}\,t \;=\; -\sum_{u=0}^{m} \frac{(T_0^k)^{m-u}}{(m-u)!} \int_0^{2^{1-k}2^N C} \frac{t^u}{u!} W(t + T_0^{k-1})\,\mathrm{d}\,t\,,$$

(taking into account that previous steps will have eliminated $\int_0^{T_0^{k-1}} \frac{t^m}{m!} W(t)\,\mathrm{d}\,t$) or equivalently (Lemmas 5.11, 5.12)

$$F_m(0; t_1^k, \ldots, t_N^k, \widetilde{T}_{N+1}, \ldots, \widetilde{T}_{r(N)})$$
$$= -\left(\frac{1}{2^{1-k}C}\right)^{m+1} \int_0^{2^{1-k}2^N C} \frac{t^m}{m!} W(t + T_0^{k-1})\,\mathrm{d}\,t$$
$$= -\int_0^{2^N C} \frac{t^m}{m!} W(2^{1-k}t + T_0^{k-1})\,\mathrm{d}\,t\,,$$

for $m = 0, 1, \ldots, N-1$.

We set $T_m^k = T_0^k + 2^{-k}Ct_m^k$ for $m = 1, \ldots, N$, and $T_m^k = T_0^k + 2^{-k}C\widetilde{T}_m$ for $m = N+1, \ldots, r(N)$. Note that $T_{r(N)}^k = T_0^{k+1}$.

Apply the switching strategy determined by $T_0^k, T_1^k, \ldots, T_{r(N)}^k$ over the time interval $[T_0^k, T_0^{k+1})$.

The algorithm can fail at this step if *either* the integrals $\int_0^{T_0^k} \frac{t^m}{m!} W(t)\,\mathrm{d}\,t$ are too large, *or* one of the 2^N switches fails to complete in the interval allotted to it. The estimate of Lemma 5.13 allows us to obtain bounds on (a) the probability of large size of the integrals $\int_0^{T_{r(N)}^k} \frac{t^m}{m!} W\,\mathrm{d}\,t = \int_0^{T_0^{k+1}} \frac{t^m}{m!} W\,\mathrm{d}\,t$, in terms of bounds on $|W|$, and (b) the probability that a switch begun at T_{r-1}^k fails to complete by time the next switch T_r^k is due to start.

Set $Z_\alpha^{-1} = \sup_{k=0,1,\ldots} (1+k)^{1+\alpha} 2^{-k/2}$, for some fixed $\alpha > 1$, and recall from Lemma 5.10 that $\varepsilon > 0$ is the bound on initial conditions required if $\kappa\varepsilon$-perturbation switching controls are guaranteed to exist. Let D_k be the event that both

$$\left| \int_0^{2^N C} \frac{t^m}{m!} W(2^{-k}t + T_0^k)\,\mathrm{d}\,t \right| \quad \leq \quad Z_\alpha (1+k)^{1+\alpha} 2^{-k/2} \varepsilon$$

for $m = 0, \ldots, N-1$, and also

$$\begin{aligned} W(T_1^k) &= -\frac{1}{2} W(T_0^k) \\ W(T_s^k) &= -W(T_{s-1}^k) \qquad \text{for } s = 1, \ldots, r(N) . \end{aligned}$$

So long as D_0, D_1, …, D_{k-1} have been satisfied, we know $W(T_0^k) = 2^{1-k}$, and so (bearing in mind the effects of the switching strategy) D_k holds only if $W(T_s^k) = 2^{-k}$ for $s = 1, \ldots, r(N)$; moreover, the condition

$$\left| \int_0^{2^N C} \frac{t^m}{m!} W(2^{1-k} t + T_0^{k-1}) \, \mathrm{d}\, t \right| \quad \leq \quad Z_\alpha (k)^{1+\alpha} 2^{-(k-1)/2} \varepsilon \quad \leq \quad \varepsilon$$

ensures we will be able to determine the solution $(t_1^k, \ldots, t_N^k)$ in Equation (25).

Now the event D_k is contained in the union of $R(n)$ events of the form

$$\begin{aligned} \sup_{u \in [T_0^k, T_{s+1}^k]} |2^{k/2} W(u)| \quad &\leq \quad \frac{Z_\alpha \varepsilon}{(2^N C)^{m+1}/(m+1)!} (1+k)^{1+\alpha} \\ &\text{and } W \text{ makes an down-crossing from } 2^{-(k-1)} \text{ to } -2^{-k} \\ &\qquad\qquad \text{over the time interval } [T_0^k, T_{s+1}^k] \, ; \end{aligned}$$

where T_0^k, T_{s+1}^k is an interval of minimum length $2^{-k}(1 - \kappa\varepsilon)$, and the down-crossing may be replaced by a down-crossing from 2^{-k} to -2^{-k}, or an up-crossing from -2^{-k} to 2^{-k} (but this does not decrease the probability of the event concerned!).

By Brownian scaling any one such event has probability bounded above by the probability of the following event:

$$\begin{aligned} \sup_{u \in [0, (1+k)^{2+2\alpha}]} |W(u)| \quad &\leq \quad \frac{Z_\alpha \varepsilon}{\sqrt{1 - \kappa\varepsilon} (2^N C)^{m+1}/(m+1)!} (1+k)^{2+2\alpha} \\ \text{and } W \text{ makes an down-crossing from } &\frac{2^{-(k-1)/2}}{\sqrt{1-\kappa\varepsilon}} (1+k)^{1+\alpha} \text{ to } -\frac{2^{-k/2}}{\sqrt{1-\kappa\varepsilon}} (1+k)^{1+\alpha} \\ &\qquad \text{over the time interval } [0, (1+k)^{2+2\alpha}] \, . \end{aligned}$$

Consequently by Lemma 5.13 we deduce that $\sum_k \mathbb{P}[D_k] < \infty$, and moreover we can use the Markov property and the density of Brownian paths to deduce there is a positive chance $p > 0$ that $\bigcap_k F_k$ occurs. If this happens then coupling succeeds at time $\lim_{k \to \infty} T_0^k$: otherwise we can start the strategy anew. We can therefore assert, almost surely success will occur eventually. $\square$

6 Impossibility of coupling all iterated integrals

Is it possible to arrange successful coupling for *all* iterated integrals at a single stopping time ζ using some adapted control J?

Summation of the coupling statements produces a statement about Laplace transforms of the path, which allows us to demonstrate that coupling of all iterated integrals is possible only in trivial cases.

Theorem 6.1 *Suppose that the initial conditions for the iterated stochastic integrals are* feasible*, in the sense that they could have been produced by integration of a continuous path starting at some previous time (without loss of generality, time* -1*). Consider an adapted control* J *producing coupling for all iterated integrals at a stopping time* ζ*. This can be produced only if* $W = B - A$ *is actually identically zero over* $[0, \zeta]$.

Proof:
Suppose the Brownian paths and all iterated integrals couple at ζ, so $W^{(n)}(\zeta) = 0$ for all n. We show that in this case $W \equiv 0$ must hold over the interval $[0, \zeta]$.

By hypothesis, we may convert into statements about integrals over $[-1, \zeta]$ (with a suitable extension of W) using powers of t. We can write

$$\sum_{n=0}^{\infty} \beta^n \int_{-1}^{\zeta} \frac{t^n}{n!} W(t) \,\mathrm{d}\, t \quad = \quad 0\,.$$

The continuous path W is bounded over $[-1, \zeta]$, so the sum on the left-hand-side converges and moreover we can exchange integral and summation to obtain

$$\int_{-1}^{\zeta} \exp(\beta t) W(t) \,\mathrm{d}\, t \quad = \quad 0$$

for all β. By uniqueness of the Laplace transform, this holds only if $W \equiv 0$ over the interval $[0, \zeta]$ as required. $\square$

Remark 6.2 This argument is essentially non-stochastic, based only on the continuity of the path which is the difference of the two coupled processes, and so holds for *any* coupling, whether co-adapted or not.

Remark 6.3 More generally, this argument extends immediately to cover for example the case when the sequence of initial conditions $W^{(m)}(0)$ is L^2 summable (use an L^2 path over $[-1, 0]$!).

7 Conclusion

We conclude by noting that the successful coupling strategies of §3 and §5 are both in essence very simple, involving switching between synchronous ($J = 1$) and mirror ($J = -1$) coupling. It would be interesting to construct a successful coupling strategy which optimized, for example, a specific exponential moment of the coupling time; one expects there would be a whole family of such couplings parametrized by the coefficient in the exponential moment, and that the coupling strategies themselves would have some kind of geometric flavour.

The results of this paper can be viewed as introducing a new notion to coupling theory: that of an "exotic coupling", a co-adapted coupling for a diffusion (in this case real Brownian motion) which successfully couples not only the diffusion itself but also a number of path functionals of the diffusion. It is striking that exotic coupling is feasible at all; the method of proof for the general case (§5) is very suggestive for how to address more general situations. Ben Arous et al. (1995) also showed the existence of an exotic coupling for planar Brownian motion using the path functional given by the Lévy stochastic area, and it would be interesting to see how far the Ben Arous et al. (1995) result could be extended to higher dimensional Brownian motion; this would be a useful next step towards the natural bold conjecture which we now present:

Conjecture 7.1 *Hypoelliptic diffusions with smooth coefficients can be coupled co-adaptively with positive chance of success from any two starting points.*

It would of course be of great interest to obtain specific applications of these couplings, perhaps for example in Coupling from the Past constructions.

Finally we remark that Price (1996) gives some results concerning exotic coupling using single functionals of the form $\int f(t) B \, \mathrm{d}\, t$.

References

Aldous, D. J. and H. Thorisson (1993). Shift-coupling. *Stochastic Processes and Their Applications 44*(1), 1–14.

Ben Arous, G., M. Cranston, and W. S. Kendall (1995). Coupling constructions for hypoelliptic diffusions: Two examples. In M. Cranston and M. Pinsky (Eds.), *Stochastic Analysis: Summer Research Institute July 11-30, 1993.*, Volume 57, Providence, RI Providence, pp. 193–212. American Mathematical Society.

Burdzy, K. and W. S. Kendall (2000, May). Efficient Markovian couplings: examples and counterexamples. *The Annals of Applied Probability 10*(2), 362–409.

Chen, X. and W. V. Li (2003). Quadratic functionals and small ball probabilities for the m-fold integrated Brownian motion. *The Annals of Probability 31*(2), 1052–1077.

Corwin, L. J. and F. P. Greenleaf (1990). *Representations of nilpotent Lie groups and their applications. Part I*, Volume 18 of *Cambridge Studies in Advanced Mathematics*. Cambridge: Cambridge University Press.

Cranston, M. and F.-Y. Wang (2000). A condition for the equivalence of coupling and shift coupling. *The Annals of Probability 28*(4), 1666–1679.

El Karoui, N. and M. Chaleyat-Maurel (1978). Un problème de réflexion et ses applications au temps local et aux équations différentielles stochastiques sur $\mathbb{R}$, cas continu. In J. Azema and M. Yor (Eds.), *Temps Locaux*, Volume 52-53, pp. 117–144.

Goldstein, S. (1978 – 1979). Maximal coupling. *Zeitschrift für Wahrscheinlichkeitstheorie und Verve Gebiete 46*(2), 193–204.

Griffeath, D. (1974 / 1975). A maximal coupling for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 31*, 95–106.

Groeneboom, P., G. Jongbloed, and J. A. Wellner (1999). Integrated Brownian motion, conditioned to be positive. *The Annals of Probability 27*(3), 1283–1303.

Hayes, T. P. and E. Vigoda (2003). A non-Markovian coloring for randomly sampling colorings. Technical report, Dept. Computer Science, University of Chicago. To appear in FOCS 2003. Preliminary version available. `http://people.cs.uchicago.edu/~vigoda/NonMarkovian.ps`.

Jerrum, M. (2003). *Counting, sampling and integrating: algorithms and complexity*. Lectures in Mathematics ETH Zürich. Basel: Birkhäuser Verlag.

Kaĭmanovich, V. A. (1986). Brownian motion and harmonic functions on covering manifolds. An entropic approach. *Dokl. Akad. Nauk SSSR 288*(5), 1045–1049.

Kendall, W. S. (1986). Stochastic differential geometry, a coupling property, and harmonic maps. *The Journal of the London Mathematical Society (Second Series) 33*, 554–566.

Kendall, W. S. (1994). Probability, convexity, and harmonic maps II: Smoothness via probabilistic gradient inequalities. *Journal of Functional Analysis 126*, 228–257.

Khoshnevisan, D. and Z. Shi (1998). Chung's law for integrated Brownian motion. *Transactions of the American Mathematical Society 350*(10), 4253–4264.

Kolmogorov, A. N. (1934). Zufällige bewegungen (zur theorie der Brownschen bewegung). *Annals of Mathematics 35*, 116–117.

Lachal, A. (1997). Local asymptotic classes for the successive primitives of Brownian motion. *The Annals of Probability 25*(4), 1712–1734.

Leeb, B. (1993). Harmonic functions along Brownian balls and the Liouville property for solvable Lie groups. *Math. Ann. 296*(4), 577–584.

Lyons, T. J. and D. Sullivan (1984). Function theory, random paths and covering spaces. *J. Differential Geom. 19*(2), 299–323.

McKean, H. (1963). A winding problem for a stochastic resonator driven by a white noise. *Journal of Mathematics of Kyoto University 2*, 227–235.

Price, C. J. (1996). *Zeros of Brownian polynomials and Coupling of Brownian areas*. Ph. D. thesis, Department of Statistics, University of Warwick.

Propp, J. G. and D. B. Wilson (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms 9*, 223–252.

Thorisson, H. (2000). *Coupling, stationarity, and regeneration*. New York: Springer-Verlag.