Pairwise Near-maximal Grand Coupling of Brownian Motions

Cheuk Ting Li\textsuperscript{a} and Venkat Anantharam\textsuperscript{b}

\textsuperscript{a}Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China, E-mail: cttli@ie.cuhk.edu.hk

\textsuperscript{b}Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA, E-mail: ananth@eecs.berkeley.edu

Abstract. The well-known reflection coupling gives a maximal coupling of two one-dimensional Brownian motions with different starting points. Nevertheless, the reflection coupling does not generalize to more than two Brownian motions. In this paper, we construct a coupling of all Brownian motions with all possible starting points (i.e., a grand coupling), such that the coupling for any pair of the coupled processes is close to being maximal, that is, the distribution of the coupling time of the pair approaches that of the maximal coupling as the time tends to 0 or \( \infty \), and the coupling time of the pair is always within a multiplicative factor \( 2e^2 \) from the maximal one. We also show that a grand coupling that is pairwise exactly maximal does not exist.

Résumé. Le couplage par réflexion de deux processus browniens est bien connu et ça donne un couplage maximal si on part de deux points distincts. Néanmoins ce couplage ne s'étend pas à plusieurs points de départ. Ici, nous construisons un couplage de tous les processus browniens à partir de tous les points de départ (c'est-à-dire un grand accouplement) où le couplage de chaque paire est presque maximal. Plus précisément, la distribution du temps de couplage de chaque paire approches celle du couplage maximal à temps zéro et à l'infini, et le temps du couplage de chaque paire est toujours pas plus grand que un multiple de \( 2e^2 \) de celui du couplage maximal. Nous démontrons également que un grand accouplement où le couplage de chaque paire de points de départ est maximal n'existe pas.

MSC2020 subject classifications: 60J65

Keywords: Brownian motion, reflection coupling, maximal coupling, grand coupling, Bessel process

1. Introduction

The maximal coupling of two stochastic processes \( P,Q \) is a coupling \( \{\{X_t\}_{t}, \{Y_t\}_{t}\} \) (i.e., the marginal distribution of \( \{X_t\}_{t} \) is \( P \), and that of \( \{Y_t\}_{t} \) is \( Q \)) that simultaneously maximizes the probability that the processes match after time \( s \) (i.e., \( X_t = Y_t \) for all \( t \geq s \)) for all \( s \). It was studied by Griffeath [12], Goldstein [11] and Pitman [27]. For two one-dimensional Brownian motions with different starting points, a maximal coupling can be given by the reflection coupling studied by Lindvall [23], Lindvall and Rogers [24], Hsu and Sturm [14], and Kendall [18]. Also see [3,16,17] for results on coupling functionals of Brownian motions.

While the maximal coupling of two stochastic processes exists under rather general conditions [11,31], it might not exist in the pairwise sense for more than two processes, that is, given a collection of stochastic processes \( \{P_\alpha\}_{\alpha \in A} \), there may not exist a coupling \( \{(X_{\alpha,t})_{t} \}_{\alpha \in A} \) (i.e., the marginal distribution of \( \{X_{\alpha,t}\}_{t} \) is \( P_\alpha \)) that simultaneously maximizes \( \mathbb{P}(\forall t \geq s : X_{\alpha,t} = X_{\beta,t}) \) for all \( s, \alpha, \beta \). The maximal coalescent coupling, which maximizes the probability that all processes in the collection match after time \( s \) (i.e., \( \mathbb{P}(\forall \alpha, \beta \in A, t \geq s : X_{\alpha,t} = X_{\beta,t}) \)), was studied by Connor [6]. Nevertheless, a maximal coalescent coupling, which only concerns whether the processes all agree after a certain time, may not give a maximal (or close to maximal) coupling when only the marginal distribution of a pair of processes \( \{X_{\alpha,t}\}_{t}, \{X_{\beta,t}\}_{t} \) is considered (refer to Section 2). Other related works on the coupling of more than two distributions or stochastic processes include coupling from the past [29,30], Wasserstein barycenter [1], and multi-marginal optimal transport [2,10,15,22,26]. A coupling of Markov chains with the same Markov kernel and all possible initial states (i.e., \( P_\alpha \) is the Markov chain starting at \( \alpha \) for any state \( \alpha \in A \)) is often called a grand coupling in the literature on coupling from the past and mixing times of Markov chains (e.g. [21]).

A classical example of a grand coupling of all one-dimensional Brownian motions with all possible starting points (i.e., \( P_\alpha = B\mathcal{M}(\alpha) \), the Brownian motion starting at \( \alpha \in \mathbb{R} \)) is the Brownian web studied by Arratia [4] and Tóth and
Werner [32]. The Brownian web has the property that, if we consider the marginal joint distribution of the processes with distribution \( \text{BM}(\alpha) \) and \( \text{BM}(\beta) \) (\( \alpha \neq \beta \in \mathbb{R} \)), then the processes move independently from \( \alpha \) and \( \beta \) respectively, until they couple (become equal), and then move together (the same as the Doeblin coupling for Markov chains [7], which is generally not maximal). The distribution of the coupling time between the two processes is the same as the distribution of twice the coupling time of the reflection coupling, i.e., the Brownian web has a multiplicative gap \( 2 \) from the optimum (refer to Section 2). The multiplicative gap does not vanish as the time tends to 0 or \( \infty \).

In this paper, we give a grand coupling \( \{(X_{\alpha,t})_{t} \}_{\alpha \in \mathbb{R}} \) of all one-dimensional Brownian motions with all possible starting points (i.e., \( \{X_{\alpha,t}\} \) has marginal \( P_{\alpha} = \text{BM}(\alpha) \) for \( \alpha \in \mathbb{R} \)), called the dyadic grand coupling, such that the coupling for any pair of the coupled processes is close to being maximal. Let

\[
\Upsilon_{\alpha,\beta} := \inf \{ s \geq 0 : X_{\gamma_1,t} = X_{\gamma_2,t}, \forall \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \}
\]

be the coupling time of all the Brownian motions with starting point lying in the interval \([\alpha, \beta] \), and \( \hat{\Upsilon}_{\alpha,\beta} := \inf \{ s \geq 0 : \hat{X}_{\alpha,t} = \hat{X}_{\beta,t}, t \geq s \} \) be the coupling time of the reflection coupling, where \( \{\hat{X}_{\alpha,t}\}_{t \geq 0}, \{\hat{X}_{\beta,t}\}_{t \geq 0} \) is the reflection coupling of \( \text{BM}(\alpha) \), \( \text{BM}(\beta) \) (which is a maximal coupling). Let “\( \preceq \)” denote first-order stochastic dominance between two real-valued random variables (i.e., \( Y \preceq Z \) if \( P(Y \geq t) \leq P(Z \geq t) \) for all \( t \in \mathbb{R} \)). Then the distribution of \( \Upsilon_{\alpha,\beta} \) is close to that of the optimal \( \hat{\Upsilon}_{\alpha,\beta} \) for all \( \alpha < \beta \), in the sense that \( \Upsilon_{\alpha,\beta} \succeq 2e^2 \hat{\Upsilon}_{\alpha,\beta} \), and the distribution of \( \Upsilon_{\alpha,\beta} \) tends to that of \( \hat{\Upsilon}_{\alpha,\beta} \) as the time tends to 0 or \( \infty \) (in the sense of multiplicative gap). More precisely, there exists a function \( r: \mathbb{R}_{>0} \to [1, 2e^2] \) (that does not depend on \( \alpha, \beta \)) such that \( \lim_{t \to 0} r(t) = \lim_{t \to \infty} r(t) = 1 \), and

\[
\hat{\Upsilon}_{\alpha,\beta} \leq \Upsilon_{\alpha,\beta} \leq r \left( \frac{\hat{\Upsilon}_{\alpha,\beta}}{|\alpha - \beta|^2} \right) \hat{\Upsilon}_{\alpha,\beta}
\]

for any \( \alpha < \beta \). Numerical evidence shows that the maximum multiplicative gap \( 2e^2 \) can be improved to around 1.5, and the dyadic grand coupling has a strictly smaller coupling time than the Brownian web in the sense of first-order stochastic dominance (see Figure 4). Refer to Section 2 for details.

A natural question is whether there exists a grand coupling \( \{\hat{X}_{\alpha,t}\}_{\alpha \in \mathbb{R}} \) of \( \{\text{BM}(\alpha)\}_{\alpha \in \mathbb{R}} \) such that any pair \( \{\hat{X}_{\alpha,t}\}_{\alpha \in \mathbb{R}}, \{\hat{X}_{\beta,t}\}_{\beta \in \mathbb{R}} \) is a maximal coupling. In Section 3, we show that such a pairwise maximal grand coupling does not exist. We conjecture that the dyadic grand coupling is optimal, in the sense of attainable failure probability bounds, as defined in Section 3.

### 2. Dyadic Grand Coupling of Brownian Motion

Let \( \text{BM}(\alpha) \) be the distribution of the standard one-dimensional Brownian motion starting at \( \alpha \in \mathbb{R} \) (\( P_{\alpha} = \text{BM}(\alpha) \) is a probability distribution over the space of continuous functions \( C([0, \infty), \mathbb{R}) \)) with the topology of uniform convergence over compact subsets of \([0, \infty)\)). We call a collection of stochastic processes \( \{\{X_{\alpha,t}\}_{t \geq 0}\}_{\alpha \in \mathbb{A}} \) (over a common probability space) indexed by \( \alpha \in \mathbb{A} \) (where \( \mathbb{A} \subseteq \mathbb{R} \)) a coupling of Brownian motions, if \( \{X_{\alpha,t}\}_{t \geq 0} \) follows \( \text{BM}(\alpha) \) for any \( \alpha \in \mathbb{A} \). If \( \mathbb{A} = \mathbb{R} \), then this is called a grand coupling.

To couple \( \text{BM}(\alpha), \text{BM}(\beta) \) with two different starting points \( \alpha, \beta \) (i.e., \( \mathbb{A} = \{\alpha, \beta\} \)), the reflection coupling [23, 24] \( \{\hat{X}_{\alpha,t}\}_{t \geq 0}, \{\hat{X}_{\beta,t}\}_{t \geq 0} \) is given by \( \{\hat{X}_{\alpha,t}\}_{t \geq 0} \sim \text{BM}(\alpha), T := \inf \{t \geq 0 : \hat{X}_{\alpha,t} = (\alpha + \beta)/2\} \), \( \hat{X}_{\beta,t} = \alpha + \beta - X_{\alpha,t} \) for \( t < T \), \( \hat{X}_{\beta,t} = X_{\alpha,t} \) for \( t \geq T \). The probability of failure of the reflection coupling is given by

\[
P\left( \exists t \geq \text{s.t.} \hat{X}_{\alpha,t} \neq \hat{X}_{\beta,t} \right) = \operatorname{erf} \left( \frac{|\alpha - \beta|}{2\sqrt{2s}} \right)
\]

for any \( s > 0 \), where

\[
\operatorname{erf}(\gamma) := \int_{-\gamma}^{\gamma} e^{-x^2} \frac{dx}{\sqrt{\pi}}
\]

is the error function.
Nevertheless, if we have to couple all the processes in \( \{ \text{BM}(\alpha) \}_{\alpha \in \mathbb{R}} \), it is impossible to simultaneously attain this probability of failure for all pairs of starting points, as will be shown in Section 3. The maximal coalescent coupling [6] is not useful in this setting since, for any fixed time, it is impossible for all the processes in \( \{ \{ X_{\alpha,t} \}_{t \geq 0} \}_{\alpha \in \mathbb{R}} \) (where \( \{ X_{\alpha,t} \}_{t \geq 0} \sim \text{BM}(\alpha) \)) to coalesce (become all equal) by that time with a positive probability. If we consider only the processes \( \{ \text{BM}(\alpha) \}_{\alpha \in [\gamma_1, \gamma_2]} \) with starting points in the interval \( A = [\gamma_1, \gamma_2] \), then a maximal coalescent coupling can be given simply by performing the reflection coupling between \( \{ X_{\gamma_1,t} \} \) and \( \{ X_{\alpha,t} \} \) for all \( \alpha \in (\gamma_1, \gamma_2) \) (note that in the reflection coupling, one process can be obtained deterministically from another, and thus we can express \( \{ X_{\alpha,t} \} \) as a function of \( \{ X_{\gamma_1,t} \} \sim \text{BM}(\gamma_1) \) for all \( \alpha \in (\gamma_1, \gamma_2) \)). This coupling is undesirable since the coupling time between \( \{ X_{\gamma_1,t} \} \) and \( \{ X_{\gamma_2-t,t} \} \) is the same as that between \( \{ X_{\gamma_1,t} \} \) and \( \{ X_{\gamma_1,t} \} \), despite \( \text{BM}(\gamma_2) \) being much closer to \( \text{BM}(\gamma_2 - \epsilon) \) than to \( \text{BM}(\gamma_1) \) for \( 0 < \epsilon < \gamma_2 - \gamma_1 \).

The Brownian web [4, 32] \( \{ \{ X_{\alpha,t}^{\text{BW}} \}_{t \geq 0} \}_{\alpha \in \mathbb{R}} \) (where \( \{ X_{\alpha,t}^{\text{BW}} \}_{t \geq 0} \sim \text{BM}(\alpha) \)) is a grand coupling which gives a probability of failure

\[
P \left( \exists t \geq s \text{ s.t. } X_{\alpha,t}^{\text{BW}} \neq X_{\beta,t}^{\text{BW}} \right)
= P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_1,t}^{\text{BW}} \neq X_{\gamma_2,t}^{\text{BW}} \right)
= \frac{\text{erf} \left( \frac{|\alpha - \beta|}{2 \sqrt{s}} \right)}{4},
\]

and hence the distribution of the coupling time between \( \{ X_{\alpha,t}^{\text{BW}} \} \) and \( \{ X_{\beta,t}^{\text{BW}} \} \) (the first time where \( X_{\alpha,t}^{\text{BW}} = X_{\beta,t}^{\text{BW}} \)) is the same as the distribution of twice the coupling time of the reflection coupling \( \{ \check{X}_{\alpha,t} \}, \{ \check{X}_{\beta,t} \} \). The multiplicative gap 2 does not vanish as the time tends to 0 or \( \infty \), that is, the Brownian web does not satisfy (1).

In this section, we propose a grand coupling that achieves a probability of failure close to that of the reflection coupling for all pairs of starting points.

Before we present our construction, we first review a connection between the Bessel process of dimension 3 and the Kolmogorov distribution. The Kolmogorov distribution arises in a Brownian bridge, as shown by Doob [8]: Let \( \{ B_t \}_{t \in [0,1]} \) be the Brownian bridge (Brownian motion conditioned on \( B_1 = 0 \)). Then

\[
P \left( \sup_{t \in [0,1]} B_t \leq \gamma \right) = 1 - \sum_{k=1}^{\infty} 2(-1)^{k+1} e^{-2k^2 \gamma^2},
\]

which is the cdf of the Kolmogorov distribution. We will use the property that if \( \{ Y_t \}_{t \geq 0} \sim \text{BES}^3(0) \), \( s > 0 \), then \( (\pi/2)^{1/2} \text{sup}_{t \leq s} Y_t \) follows the Kolmogorov distribution, i.e.,

\[
P \left( \text{sup}_{t \leq s} Y_t \leq \gamma \right) = \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( \frac{k^2 \pi^2 s}{2 \gamma^2} \right).
\]

See [5, 13, 19]. We now prove this property for the sake of completeness. By [5, Theorem 2],

\[
P \left( \text{sup}_{t \leq s} Y_t \leq \gamma \right) = \frac{1}{2^{-1/2} \Gamma(3/2)} \sum_{k=1}^{\infty} \frac{j_{1/2,k}^{-1/2}}{J_{3/2}(j_{1/2,k})} \exp \left( \frac{j_{1/2,k}^2 s}{2 \gamma^2} \right).
\]
where $J_\alpha(x)$ is the Bessel function of the first kind, and $j_{\alpha,k}$ is its increasing sequence of positive roots. We have $J_{1/2}(x) = \sqrt{2/\pi} \sin(x)$, $J_{3/2}(x) = \sqrt{2/\pi} (\sin(x)/x - \cos(x))$, and hence $J_{1/2,2} = \pi k$.

We give an intuitive description of the construction of the grand coupling. First consider a coupling of all Brownian motions starting at integer points, i.e., $\{BM(\alpha)\}_{\alpha \in \mathbb{Z}}$. An obvious construction is to couple every consecutive pair $BM(\alpha), BM(\alpha + 1)$ using reflection coupling. Note that in the reflection coupling of $BM(\alpha), BM(\alpha + 1)$, the Brownian motion following $BM(\alpha + 1)$ can be expressed as a function of the Brownian motion following $BM(\alpha)$, and vice versa. Therefore, once we generate the Brownian motion following $BM(0)$, all the other Brownian motions in the coupling can be obtained as functions of that Brownian motion. We denote this coupling as $\{\hat{X}_{\alpha,t}^{[0]}\}_{\alpha \in \mathbb{Z}, t \geq 0}$, where $\{\hat{X}_{\alpha,0}^{[0]}\}_{t \geq 0} \sim BM(\alpha)$.

An informal way to construct the grand coupling of Brownian motions starting at any real number is to consider $\{\epsilon_\alpha \hat{X}_{\alpha,t}^{[0]}\}_{\alpha \in \mathbb{Z}, t \geq 0}$, which is a coupling of Brownian motions starting at multiples of $\epsilon$ (where $\epsilon > 0$), and take the “limit” as $\epsilon \to 0$. The problem is that $\epsilon_\alpha \hat{X}_{\alpha,t}^{[0]}$ does not converge almost surely as $\epsilon \to 0$ (though it might be possible to argue that the joint distribution of $\{\epsilon_\alpha \hat{X}_{\alpha,t}^{[0]}\}_{\alpha \in \mathbb{Z}, t \geq 0}$ converges in some sense). Instead, we will present a more constructive way to obtain the coupling.

To construct the coupling, we modify the coupling $\{\hat{X}_{\alpha,t}^{[0]}\}_{\alpha \in \mathbb{Z}, t \geq 0}$ to Brownian motions starting at half integers, i.e., $\{BM(\alpha)\}_{\alpha \in \mathbb{Z}/2}$ (we write $\mathbb{Z}/2 = \{k/2 : k \in \mathbb{Z}\}$), while retaining the previous coupling $\{\hat{X}_{\alpha,0}^{[0]}\}$ as much as possible. To accomplish this, we first construct a coupling $\{Z_{\alpha,t}\}_{\alpha \in \mathbb{Z}/2, t \geq 0}$, where $\{Z_{\alpha,t}\}_{t \geq 0} \sim BM(\alpha)$, and $\{Z_{\alpha,t}\}_{t \geq 0}, \{Z_{\alpha+1/2,t}\}_{t \geq 0}$ follows the reflection coupling for any $\alpha \in \mathbb{Z}/2$, i.e., we repeat the aforementioned construction with reflection coupling between consecutive pairs. Nevertheless, we will only use $\{Z_{\alpha,0}\}$ up to the first coalescent time where any consecutive pair coalesces (we use “coalescent time” and “coalescing time” interchangeably hereafter). Let $T := \inf\{t \geq 0 : Z_{0,t} = Z_{1/2,t} \text{ or } Z_{0,t} = Z_{-1/2,t}\}$, $W := 1$ if $Z_{0,T} = Z_{1/2,T}$, $W := -1$ if $Z_{0,T} = Z_{-1/2,T}$. Note that if the pair of processes $\{Z_{0,t}\}, \{Z_{1/2,t}\}$ is the first to coalesce, then the pair $\{Z_{1,t}\}, \{Z_{3/2,t}\}$ coalesces at the same time (since $\{Z_{1,t}\} = \{Z_{0,t}\}$ reflected twice before the coalescent time, it is simply $\{Z_{0,t}\}$ shifted by 1 before the coalescent time), and in general each of the pairs $\{Z_{k,t}\}, \{Z_{k+1/2,t}\}$ for $k \in \mathbb{Z}$ coalesces at the same time. Also, the pair of processes $\{Z_{0,t}\}, \{Z_{-1/2,t}\}$ is the first to coalesce, then the pairs $\{Z_{k,t}\}, \{Z_{k-1/2,t}\}$ for $k \in \mathbb{Z}$ coalesce at the same time. Therefore, if any pair coalesces, then either all pairs $\{Z_{k,t}\}, \{Z_{k+1/2,t}\}$ for $k \in \mathbb{Z}$ coalesce ($W = 1$), or the set of positions at the coalescent time is $\{Z_{0,T} : \alpha \in \mathbb{Z}/2\} = \mathbb{Z} + 1/4$, or all pairs $\{Z_{k,t}\}, \{Z_{k-1/2,t}\}$ for $k \in \mathbb{Z}$ coalesce ($W = -1$), and $\{Z_{0,T} : \alpha \in \mathbb{Z}/2\} = \mathbb{Z} - 1/4$. We then simply weld $\{Z_{\alpha,t}\}_{\alpha \in \mathbb{Z}/2, t \leq T}$ together with $\{\hat{X}_{\alpha,0}^{[0]}\}_{t \geq 0}$ shifted by either 1/4 or -1/4. More precisely, we let

$$
\hat{X}_{\alpha,0}^{[0]} := \begin{cases} 
Z_{\alpha,t} & \text{if } t \leq T \\
\hat{X}_{\alpha+1/2-W/4,t-T}^{[0]} + W/4 & \text{if } t > T 
\end{cases}
$$

for $\alpha \in \mathbb{Z}/2, t \geq 0$. We have $\{\hat{X}_{\alpha,0}^{[-1]}\}_{t \geq 0} \sim BM(\alpha)$ for $\alpha \in \mathbb{Z}/2$. Refer to Figure 1 for an illustration of the construction.

Alternatively, we can first generate $W \sim \text{Unif}\{\pm 1\}$, and then generate $\{Z_{\alpha,t}\}$ conditional on $W$. Conditional on $W = 1$, we know $\{Z_{0,t}\}$ hits 1/4 (so $\{Z_{0,t}\}, \{Z_{1/2,t}\}$ coalesce) before −1/4. Therefore, the conditional distribution of $\{Z_{0,t}\}$ given $W = 1$ is $\text{BES}^3(1/4) - 1/4$, a Bessel process of dimension 3 starting at 1/4, shifted by −1/4, stopped when it hits 1/4. Considering both cases for $W$, the conditional distribution of $\{Z_{0,t}\}$ given $W$ is $W \cdot (\text{BES}^3(1/4) - 1/4)$.

We continue this procedure to define $\hat{X}_{\alpha,t}^{[-2]}, \hat{X}_{\alpha,t}^{[-3]} \ldots$ We then obtain a sequence $\{W_j\}_{j \geq 0} \sim \text{Unif}\{\pm 1\}$ and a sequence of stopped shifted Bessel processes. Now we construct a “limit” $\hat{X}_{\alpha,t}^{[-\infty]}$ of $\hat{X}_{\alpha,t}^{[j]}$ as $j \to -\infty$. Note that the stopped shifted Bessel process (ignoring the sign $W_j$) in the definition of $\hat{X}_{\alpha,t}^{[j]}$ (e.g. $\text{BES}^3(1/4) - 1/4$ for $j = -1$) starts at 0 and stops when it hits 21−1. We shift it so it starts at 21−1 and stops at 21, and weld these for all $j$ together to form an (unshifted) Bessel process starting at 0. Let this Bessel process be $Y_j$. To compute $\hat{X}_{\alpha,t}^{[-\infty]}$, we have to add back the signs $W_j$. It can be observed in (4) that when $\alpha = 0$, $\alpha + 1/2 - W/4 = 0$. Hence, $\hat{X}_{0,t}^{[j]}$ is a shifted version of $\hat{X}_{0,t}^{[j+1]}$ after the coalescent time. Therefore, to find $\hat{X}_{0,t}^{[-\infty]}$, we only need to take the limit of $\hat{X}_{0,t}^{[j]}$. Let $T_j := \inf\{t \geq 0 : Y_t = 2^j\}$. Then
Pairwise Near-maximal Grand Coupling

For $t \in [T_{j-1}, T_j]$ corresponds to $\hat{X}_{0,t}^{[j]}$, and this segment of $Y_t$ should be flipped according to the sign $W_j$. Hence, we obtain

$$\hat{X}_{0,t}^{[-\infty]} = \sum_{k=-\infty}^{j-1} W_k 2^{k-1} + (Y_t - 2^{j-1})W_j$$

for $t \in [T_{j-1}, T_j]$. The term $\sum_{k=-\infty}^{j-1} W_k 2^{k-1}$ is to make the process continuous in $t$ and start at 0.

To obtain $\hat{X}_{\alpha,t}^{[-\infty]}$ for general $\alpha$, since we no longer have the property that $|\alpha + 1/2 - W/4| = \alpha$ as when $\alpha = 0$ in (4), the signs to multiply the shifted version of $Y_t$ with depend on both $W_j$ and $\alpha$ (we will see in (6) that the signs depend on the binary representation of $\alpha$). Intuitively, for each $j$, we decide whether to flip the segment $\{Y_t\}_{t \in [0,T_j]}$ of the Bessel process $Y_t$ along the line $y = Y_{T_j} = 2^j$. Without any flipping, the process $Y_t$ starts at 0. With only one flipping along $y = Y_{T_j} = 2^j$, the resultant process starts at $2^{j+1}$. We can make the process start at any nonnegative point $\alpha$ by choosing which $j$ to flip according to the binary representation of $\alpha$.

The actual situation is more complicated since the resultant process should be a Brownian motion instead of a Bessel process. The random signs $W_j$ are added for this purpose. If we flip the Bessel process using the signs $W_j$ according to (5), the resultant process will be a Brownian motion starting at 0. It can be observed from (5) that if we want to flip $W_j$ while keeping the tail of $\hat{X}_{0,t}^{[-\infty]}$ unchanged, we have to shift the starting point of the process by $2^j W_j$ (since the coefficient of $W_j$ is $2^{j-1}$). Therefore, we can choose which $W_j$ to flip according to the binary representation of $\alpha$ (or, more precisely, the representation of $\alpha$ in the mixed radix numeral system where position $j$ has value $2^j W_j$) to make the process start at $\alpha$.

Let $G_{\alpha,j}$ (instead of $W_j$) be the sign applied to the shifted version of $\{Y_t\}_{t \in [T_{j-1}, T_j]}$ for $\hat{X}_{\alpha,t}^{[-\infty]}$, i.e.,

$$\hat{X}_{\alpha,t}^{[-\infty]} = \alpha + \sum_{k=-\infty}^{j-1} G_{\alpha,k} 2^{k-1} + (Y_t - 2^{j-1})G_{\alpha,j}$$

for $t \in [T_{j-1}, T_j]$. Intuitively, $\{G_{\alpha,j}\}_j$ is the sequence of flippings that makes the process start at $\alpha$. We have

$$\hat{X}_{\alpha,t}^{[-\infty]} - \hat{X}_{0,t}^{[-\infty]} = \alpha + \sum_{k=-\infty}^{j-1} (G_{\alpha,k} - W_k) 2^{k-1} + (Y_t - 2^{j-1})(G_{\alpha,j} - W_j).$$
Since $\hat{X}_{\alpha,t}^{[-\infty]}$ should coalesce with $\hat{X}_{0,t}^{[-\infty]}$ for large $t$, we have

$$\sum_{k=-\infty}^{\infty} (W_k - G_{\alpha,k})2^{k-1} = \alpha. \quad (6)$$

This can be regarded as the aforementioned mixed radix numerical system representation of $\alpha$ since each $k$ where $G_{\alpha,k} \not= W_k$ increases the left hand side by $2^k W_k$. This formula can be used to obtain $G_{\alpha,j}$ from the binary representation of $\alpha$.

Note that the above construction is “nonuniform” since we flip $Y_i$ only at the points $T_j$ where $Y_{T_j} = 2^j$ (i.e., we favor powers of 2). Such nonuniformity causes the coupling time between $\hat{X}_{\alpha,t}^{[-\infty]}$ and $\hat{X}_{\beta,t}^{[-\infty]}$ to depend not only on $|\alpha - \beta|$, but also on the values of $\alpha$ and $\beta$ (and how they align with the grid points of step size $2^j$). To make the construction uniform, we add a “dithering” term, which is a random number $\theta \in [0, 1]$ to the exponent (it becomes $2^{j+\theta}$) so the flipping points are uniform in the logarithmic scale. The $G_{\alpha,j}$ would become $G_{\theta,\alpha,j}$.

The explicit construction of the grand coupling $\{\{X_{\alpha,t} \}_{t \geq 0}\}_{\alpha \in \mathbb{R}}$ (where $\{X_{\alpha,t} \}_{t \geq 0} \sim \text{BM}(\alpha)$) is given in the following definition.

**Definition 2.1** (Dyadic grand coupling of Brownian motion). Let $\{Y_t \}_{t \geq 0} \sim \text{BES}^3(0)$, the Bessel process of dimension 3 starting at 0. Let $W_i \sim \text{Unif}\{\pm 1\}$ for $i \in \mathbb{Z}$ be independent of $\{Y_t \}_{t \geq 0}$. For any $\theta \in [0, 1]$ and $\alpha \in \mathbb{R}$, let

$$G_{\theta,\alpha,j} = \begin{cases} W_j & \text{if } \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^{j+\theta-1} \right) \mod 2^{j+\theta+1} \in [0, 2^{j+\theta}) \\ -W_j & \text{otherwise,} \end{cases}$$

for $j \in \mathbb{Z}$, where $a \mod b := a - b\lfloor a/b \rfloor$ for $b > 0$. While this formula may look complicated, it is merely the mixed radix numeral system representation of $\alpha$ where position $k$ has value $2^k W_k$ (the digit at position $k$ is 1 if $G_{\theta,\alpha,k} \not= W_k$, and the digit is 0 otherwise).

By the definition of $G_{\theta,\alpha,j}$, the conditional distribution of $G_{\theta,\alpha,j}$ given any $\{W_k \}_{k < j}$ is $\text{Unif}\{\pm 1\}$ (since $W_j \sim \text{Unif}\{\pm 1\}$ independent of $\{W_k \}_{k < j}$). Hence $G_{\theta,\alpha,j} \sim \text{Unif}\{\pm 1\}$, and is independent of $\{Y_t \}_{t \geq 0}$.

As will be shown in Appendix A, if $\sup \{j : W_j = 1\} = \sup \{j : W_j = -1\} = \infty$ (which happens almost surely), then

$$\left[ 2^{-(j+\theta)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} \right) + \frac{1}{2} \right] = \sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k})2^{k-j-1} \quad (7)$$

for any $j \in \mathbb{Z}$,

$$\sum_{j=-\infty}^{\infty} (W_j - G_{\theta,\alpha,j})2^{j+\theta-1} = \alpha, \quad (8)$$

and for any fixed $\theta, \alpha$, we have $G_{\theta,\alpha,j} = W_j$ for all sufficiently large $j$. For $\theta \in [0, 1]$ and $i \in \mathbb{Z}$, let

$$T_{\theta,i} := \inf \{t \geq 0 : Y_t = 2^{i+\theta} \}.$$

Let $X_{\theta,\alpha,0} := \alpha$. For $t > 0$, with $i$ defined to satisfy $T_{\theta,i-1} < t \leq T_{\theta,i}$, let

$$X_{\theta,\alpha,t} := \alpha + \sum_{j=-\infty}^{i-1} G_{\theta,\alpha,j}2^{j+\theta-1} + (Y_t - 2^{i+\theta-1})G_{\theta,\alpha,i} \quad (9)$$

$$= (Y_t - 2^{i+\theta})G_{\theta,\alpha,i} + \sum_{j=-\infty}^{\infty} (W_j - 1_j \{j > i\} G_{\theta,\alpha,j})2^{j+\theta-1}, \quad (10)$$

where the equivalence of (9) and (10) can be seen by (8). Let $\Theta \sim \text{Unif}[0, 1]$ be independent of $\{(Y_t)_{t \geq 0}, \{W_i\} \}$. Define the coupling as

$$X_{\alpha,t} := X_{\Theta,\alpha,t}.$$
We now check that \( \{X_{\alpha,t}\}_{t \geq 0} \sim \text{BM}(\alpha) \). First, for any \( i \), we can see from (9) that \( \lim_{t \to \tau_{i,\alpha,T}^{-}} X_{\theta,\alpha,t} = X_{\theta,\alpha,T_{i,\alpha,-1}} \), and hence \( \{X_{\alpha,t}\}_{t \geq 0} \) is continuous in \( t \).

By the strong Markov property, \( \{Y_{t+T_{m}}\}_{t \geq 0} \sim \text{BES}^3(2^{i+\theta-1}) \). Since \( \text{BES}^3(2^{i+\theta-1}) \) is the distribution of \( \text{BM}(2^{i+\theta-1}) \) conditioned to stay positive [20, 25, 28, 33], we see that \( \{Y_{t+T_{m}}\}_{0 \leq t \leq T_{i,\alpha,-1}} \) has the same distribution as \( \text{BM}(2^{i+\theta-1}) \) conditioned to stay positive and stopped when it hits \( 2^{i+\theta} \), or equivalently, stopped when it hits either 0 or \( 2^{i+\theta} \) and conditioned on the event that it hits \( 2^{i+\theta} \) (which has probability 1/2, so the conditioning is well-defined). By symmetry, \( \{Y_{t+T_{m}} - 2^{i+\theta-1}G_{\theta,\alpha,i}\}_{0 \leq t \leq T_{i,\alpha,-1}} \) has the same distribution as \( \text{BM}(0) \) stopped when it hits either \( 2^{i+\theta-1} \) or \( -2^{i+\theta-1} \), and is independent of \( \{T_{i,\alpha,-1}, Y_{t}\}_{t \leq T_{i,\alpha,-1}} \) (since \( G_{\theta,\alpha,i} \sim \text{Unif} \{\pm 1\} \) independent of \( \{Y_{t}\}_{t \geq 0} \)). Welding these processes together, we can see from (9) that \( \{X_{\theta,\alpha,t+T_{m}} - X_{\theta,\alpha,T_{m}}\}_{t \geq 0} \) follows \( \text{BM}(0) \) and is independent of \( \{T_{i,\alpha,-1}, Y_{t}\}_{t \leq T_{i,\alpha,-1}} \) for any \( \theta, \alpha, i \). Since a random process with continuous sample paths is characterized by its finite-dimensional marginals, by letting \( i \to -\infty \), we see that \( \{X_{\theta,\alpha,t}\}_{t \geq 0} \sim \text{BM}(\alpha) \) for any \( \theta, \alpha \). More precisely, for any \( 0 < \tau_{1} < \cdots < \tau_{m} \), we have \( \{X_{\theta,\alpha,T_{\gamma_{k}},t} - X_{\theta,\alpha,T_{\gamma_{k}}\tau_{k}}\}_{k \to \tau_{k}} \to \text{BM}(\alpha) \) as \( i \to -\infty \) almost surely since \( T_{\gamma_{k}} \to 0 \) and \( X_{\theta,\alpha,T_{\gamma_{k}}} \to \alpha \). Since \( \{X_{\theta,\alpha,T_{\gamma_{k}},t} - X_{\theta,\alpha,T_{\gamma_{k}}}\}_{k} \) has the same distribution as \( \{B_{\gamma_{k}}\}_{k} \) (where \( \{B_{t}\}_{t} \) is the Brownian motion), \( \{X_{\theta,\alpha,T_{\gamma_{k}}}, \alpha\}_{k} \) also has the same distribution as \( \{B_{\gamma_{k}}\}_{k} \). Hence \( \{X_{\alpha,t}\}_{t \geq 0} \sim \text{BM}(\alpha) \).

We then evaluate the probability of failure of this coupling.

**Theorem 2.2.** *For the dyadic grand coupling of Brownian motion \( \{\{X_{\alpha,t}\}_{t \geq 0}\}_{\alpha \in \mathbb{R}} \), we have

\[
\mathbb{P} \left( \exists t \geq s \text{ s.t. } X_{\alpha,t} \neq X_{\beta,t} \right)
= \mathbb{P} \left( \exists \gamma_{1}, \gamma_{2} \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_{1},t} \neq X_{\gamma_{2},t} \right)
= \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^{2}\pi^{2}}{2s^{2}} \right) \zeta^{-2}\psi \ln 2 \right) \left( \min \{\zeta\psi^{-1}, 1\} - \frac{1}{2} \right) d\zeta
\]

for any \( \alpha < \beta \) and \( s > 0 \), where \( \psi := |\alpha - \beta|/\sqrt{5} \).

As a consequence, we have the following results.

**Corollary 2.3.** *Let \( \{\{X_{\alpha,t}\}_{t \geq 0}\}_{\alpha \in \mathbb{R}} \) be the one-dimensional dyadic grand coupling of Brownian motion. Fix any \( \alpha < \beta \). Let \( \{X_{\alpha,t}, \{\hat{X}_{\beta,t}\}_{t \geq 0} \) be the reflection coupling of \( \text{BM}(\alpha), \text{BM}(\beta) \). Let \( T_{\alpha,\beta} := \inf \{s \geq 0 : X_{\gamma_{1},t} = X_{\gamma_{2},t} \} \) for any \( \gamma_{1}, \gamma_{2} \in [\alpha, \beta] \) satisfying \( \gamma_{1} \neq \gamma_{2} \) and \( t \geq s \). Theorem 2.2 implies that

\[
\mathbb{P} \left( \exists t \geq s \text{ s.t. } X_{\alpha,t} \neq X_{\beta,t} \right)
= \mathbb{P} \left( \exists \gamma_{1}, \gamma_{2} \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_{1},t} \neq X_{\gamma_{2},t} \right)
= \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^{2}\pi^{2}}{2s^{2}} \right) \zeta^{-2}\psi \ln 2 \right) \left( \min \{\zeta\psi^{-1}, 1\} - \frac{1}{2} \right) d\zeta
\]

for any \( \alpha < \beta \) and \( s > 0 \), where \( \psi := |\alpha - \beta|/\sqrt{5} \).
$X_{\gamma_1,t}, \forall \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s$ and $\mathbf{\hat{Y}}_{\alpha,\beta} := \inf\{s \geq 0 : \mathbf{\hat{X}}_{\alpha,t} = \mathbf{\hat{X}}_{\beta,t}, \forall t \geq s\}$ be the coupling times of the dyadic grand coupling and the reflection coupling respectively. Let $F_{\mathbf{\hat{Y}}_{\alpha,\beta}}^{-1}(p) := \inf\{s : P(\mathbf{\hat{Y}}_{\alpha,\beta} \leq s) \geq p\}$ be the inverse distribution function of $\mathbf{\hat{Y}}_{\alpha,\beta}$, and define $F_{\mathbf{Y}_{\alpha,\beta}}^{-1}(p)$ similarly. We have:

1. For any $s > 0$ (let $\psi := |\alpha - \beta|/\sqrt{s}$),
   $$P(\mathbf{Y}_{\alpha,\beta} > s) \leq \frac{\psi}{\sqrt{2\pi}}.$$

   As a result,
   $$\lim_{s \to \infty} \frac{P(\mathbf{Y}_{\alpha,\beta} > s)}{P(\mathbf{\hat{Y}}_{\alpha,\beta} > s)} = 1,$$

   and
   $$\lim_{p \to 1} \frac{F_{\mathbf{Y}_{\alpha,\beta}}^{-1}(p)}{F_{\mathbf{\hat{Y}}_{\alpha,\beta}}^{-1}(p)} = 1,$$

   i.e., the tail of the distribution of the coupling time of the dyadic grand coupling approaches that of the reflection coupling as $s \to \infty$.

2. For any $s > 0$ (let $\psi := |\alpha - \beta|/\sqrt{s}$), if $\psi \geq 2\sqrt{2}$, then
   $$P(\mathbf{Y}_{\alpha,\beta} > s) \leq 1 - \left(1 - \left(\text{erf}\left(\frac{\psi + 8/\psi}{2\sqrt{2}}\right)\right)^3\right) \ln(1+8/\psi^2) + (1+8/\psi^2)^{-1} - 1.$$

   As a result,
   $$\lim_{p \to 0} \frac{F_{\mathbf{Y}_{\alpha,\beta}}^{-1}(p)}{F_{\mathbf{\hat{Y}}_{\alpha,\beta}}^{-1}(p)} = 1,$$

   i.e., the multiplicative gap between $\mathbf{Y}_{\alpha,\beta}$ and $\mathbf{\hat{Y}}_{\alpha,\beta}$ vanishes as $s \to 0$.

3. For any $s > 0$ (let $\psi := |\alpha - \beta|/\sqrt{s}$),
   $$P(\mathbf{Y}_{\alpha,\beta} > s) \leq \text{erf}\left(\frac{e\psi}{2}\right).$$

As a result, $2e^2\mathbf{\hat{Y}}_{\alpha,\beta}$ first-order stochastically dominates $\mathbf{Y}_{\alpha,\beta}$, i.e., the dyadic grand coupling is pairwise within a multiplicative factor $2e^2$ from being maximal.

These three bounds imply (1) by taking
$$r(t) := \frac{1}{t} F_{\mathbf{Y}_{0,1}}^{-1}(F_{\mathbf{\hat{Y}}_{0,1}}(t)).$$

Note that $\mathbf{Y}_{\alpha,\beta}/|\alpha - \beta|^2$ has the same distribution as $\mathbf{Y}_{0,1}$, and $\mathbf{\hat{Y}}_{\alpha,\beta}/|\alpha - \beta|^2$ has the same distribution as $\mathbf{\hat{Y}}_{0,1}$.

We first prove Theorem 2.2.

**Proof of Theorem 2.2.** Intuitively, $X_{\alpha,t}$ depends on the mixed radix numeral system representation of $\alpha$ where position $k$ has value $2^{k+\theta}w_k$. Therefore, the coupling time of $X_{\alpha,t}$ and $X_{\beta,t}$ is related to the number of digits that the representations of $\alpha$ and $\beta$ agree on (the more number of digits that they agree on, the earlier the coupling time is). Recall that $X_{\alpha,t}$ is defined piecewise on intervals $[T_{\theta,i-1}, T_{\theta,i}]$. Let $I$ be such that $T_{\theta,I-1} < s \leq T_{\theta,I}$. Intuitively, the time $s$ corresponds to the position $I$ in the mixed radix numeral system representation, and if the representations of $\alpha$ and $\beta$ agree up to position $I$, then $X_{\alpha,t}$ and $X_{\beta,t}$ couple before time $s$. We have
$$P\left(\exists t \geq s \text{ s.t. } X_{\alpha,t} \neq X_{\beta,t}\right)$$
Fig 3. Plot of the cumulative distribution function $F_{\bar{\Upsilon}_{0,1}}$ (black), $F_{\hat{\Upsilon}_{0,1}}$ (blue), the bound on $F_{\Upsilon_{0,1}}$ in Corollary 2.3 (red) (we take the pointwise maximum of the three bounds in Corollary 2.3), and the cumulative distribution function of the coupling time of the Brownian web (dashed line). The left figure is in log-scale for the x-axis, whereas the right figure is in log-scale for both axes. Note that $F_{\Upsilon_{0,1}}$ (the black curve) is bounded between the blue curve and the red curve. Due to numerical precision issues, $F_{\Upsilon_{0,1}}(t)$ is not plotted for small $t$’s in the right figure.

Fig 4. Plot of $F_{\bar{\Upsilon}_{\alpha,\beta}}^{-1}(p)/F_{\hat{\Upsilon}_{\alpha,\beta}}^{-1}(p)$ (black), the upper bound on $F_{\bar{\Upsilon}_{\alpha,\beta}}^{-1}(p)/F_{\hat{\Upsilon}_{\alpha,\beta}}^{-1}(p)$ in Corollary 2.3 (red), and the corresponding ratio for the Brownian web (dashed line, which is constantly 2) against $p$, where $\alpha < \beta$ (these curves do not depend on the choice of $\alpha, \beta$). While Corollary 2.3 gives a multiplicative gap $2e^2$, we can observe in this graph that the multiplicative gap can be improved to around 1.5, since $F_{\bar{\Upsilon}_{\alpha,\beta}}^{-1}(p)/F_{\hat{\Upsilon}_{\alpha,\beta}}^{-1}(p)$ stays below 1.5 for all $p$.

\[
= P \left( \exists k \geq I \text{ s.t. } G_{\Theta,\alpha,k} \neq G_{\Theta,\beta,k} \right) \\
= P \left( \exists k \geq I \text{ s.t. } \sum_{j=k}^{\infty} (W_j - G_{\Theta,\alpha,j})2^{j-k-1} \neq \sum_{j=k}^{\infty} (W_j - G_{\Theta,\beta,j})2^{j-k-1} \right),
\]

where the first equality comes from (10) (since $X_{\alpha,t}$ depends on $\alpha$ only through $\{G_{\Theta,\alpha,k}\}_{k \geq t}$), and the second equality is because there is a one-to-one correspondence between $\{G_{\Theta,\alpha,j}\}_{j \geq k}$ and $\sum_{j=k}^{\infty} (W_j - G_{\Theta,\alpha,j})2^{j-k-1}$ when $\{W_j\}$ is given. Loosely speaking, the last line is the probability that the mixed radix numeral system representations of $\alpha$ and $\beta$ do not agree up to position $I$. By (7), we have

\[
P \left( \exists t \geq s \text{ s.t. } X_{\alpha,t} \neq X_{\beta,t} \right)
\]
\[ P \left( \exists k \geq I \text{ s.t. } \sum_{j=k}^{\infty} (W_j - G_{\Theta, \alpha, j}) 2^{j-k-1} \neq \sum_{j=k}^{\infty} (W_j - G_{\Theta, \beta, j}) 2^{j-k-1} \right) \]

\[ = P \left( \exists k \geq I \text{ s.t. } \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \right) \]

\[ = P \left( s \leq T_{\Theta, \max} \{ k : \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \} \right) \]

\[ = P \left( \sup_{t \leq s} Y_t \leq 2^{\max \{ k : \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \} + \theta \right) \]

\[ = E \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^2 \pi^2}{2Z^2} \right) \right), \]

where

\[ Z := s^{-1/2} 2^{\max \{ k : \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \} + \theta, \]

and the last equality is because \((\pi/2)^{1/2} / \sup_{t \leq s} Y_t \) follows the Kolmogorov distribution (3). By the same arguments,

\[ P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \]

\[ = P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], k \geq I \text{ s.t. } G_{\Theta, \gamma_1, k} \neq G_{\Theta, \gamma_2, k} \right) \]

\[ = P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], k \geq I \text{ s.t. } \left| 2^{-(k+\theta)} \left( \gamma_1 - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \gamma_2 - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \right) \]

\[ = P \left( \exists k \geq I \text{ s.t. } \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \right), \]

and hence the two probabilities in Theorem 2.2 are equal.

It is left to find the distribution of \( Z \), which depends on the number of digits that the mixed radix numeral system representations of \( \alpha \) and \( \beta \) agree on. To check whether the binary representations of \( \alpha \) and \( \beta \) agree up to position \( l \), it suffices to check whether \( [2^{-l} \alpha] = [2^{-l} \beta] \) (which is equivalent to \( \forall k \geq l : [2^{-k} \alpha] = [2^{-k} \beta] \)). A similar result applies to the mixed radix numeral system. By (7), the probability that the mixed radix numeral system representations do not agree up to position \( l \) is

\[ P \left( \max \left\{ k : \left| 2^{-(k+\theta)} \left( \alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(k+\theta)} \left( \beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \right\} \geq l \right) \]

\[ = P \left( \exists k \geq l \text{ s.t. } \sum_{j=k}^{\infty} (W_j - G_{\Theta, \alpha, j}) 2^{j-k-1} \neq \sum_{j=k}^{\infty} (W_j - G_{\Theta, \beta, j}) 2^{j-k-1} \right) \]

\[ = P \left( \sum_{j=l}^{\infty} (W_j - G_{\Theta, \alpha, j}) 2^{j-l-1} \neq \sum_{j=l}^{\infty} (W_j - G_{\Theta, \beta, j}) 2^{j-l-1} \right) \]

\[ = P \left( \left| 2^{-(l+\theta)} \left( \alpha - \sum_{j=-\infty}^{l-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \neq \left| 2^{-(l+\theta)} \left( \beta - \sum_{j=-\infty}^{l-1} W_j 2^{j+\theta-1} \right) + \frac{1}{2} \right| \right) \]
\[
\begin{align*}
&= \int_0^1 1 \left\{ \left[ 2^{-(l+\theta)} (\alpha - (1 - 2\gamma)2^{l+\theta-1}) + \frac{1}{2} \right] \neq \left[ 2^{-(l+\theta)} (\beta - (1 - 2\gamma)2^{l+\theta-1}) + \frac{1}{2} \right] \right\} d\gamma \\
&= \int_0^1 1 \left\{ \left[ 2^{-(l+\theta)} \alpha + \gamma \right] \neq \left[ 2^{-(l+\theta)} \beta + \gamma \right] \right\} d\gamma \\
&= \min \left\{ 2^{-(l+\theta)} |\alpha - \beta|, 1 \right\}.
\end{align*}
\]

Combining (12) and (13), we have

\[
P(Z \geq \zeta) = P \left( \max \left\{ k : \left[ 2^{-(k+\theta)} (\alpha - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1}) + \frac{1}{2} \right] \neq \left[ 2^{-(k+\theta)} (\beta - \sum_{j=-\infty}^{k-1} W_j 2^{j+\theta-1}) + \frac{1}{2} \right] \right\} \geq \left\lfloor \log_2 (\zeta/\sqrt{s}) - \Theta \right\rfloor \right)
\]

\[
= \int_0^1 \min \left\{ 2^{-(\lfloor \log_2 (\zeta/\sqrt{s}) - \Theta \rfloor + \theta)} |\alpha - \beta|, 1 \right\} d\theta
\]

\[
= \int_0^1 \min \left\{ \zeta^{-1} s^{-1/2} 2^{-\theta} |\alpha - \beta|, 1 \right\} d\theta
\]

\[
= \int_0^1 \min \left\{ \zeta^{-1} 2^{-\theta} \psi, 1 \right\} d\theta,
\]

where \( \psi := |\alpha - \beta|/\sqrt{s} \), and (a) is because \([x - \Theta] + \Theta\) has the same distribution as \(x + \Theta\) for any fixed \(x\), where \(\Theta \sim \text{Unif}[0, 1]\) (since \([x - \Theta] - (x - \Theta) \sim \text{Unif}[0, 1]\)). Hence,

\[
- \frac{d}{d\zeta} P(Z \geq \zeta)
\]

\[
= - \int_0^1 \frac{d}{d\zeta} \min \left\{ \zeta^{-1} 2^{-\theta} \psi, 1 \right\} d\theta
\]

\[
= \int_0^1 1 \left\{ \zeta^{-1} 2^{-\theta} \psi \leq 1 \right\} \zeta^{-2} 2^{-\theta} \psi d\theta
\]

\[
= \zeta^{-2} \psi \int_0^1 \frac{2^{-\theta} d\theta}{\min(\log_2(\zeta^{-1} \psi), 0, 1)} - \frac{1}{2}
\]

\[
= \frac{\zeta^{-2} \psi}{\ln 2} \left( 2^{-\min(\log_2(\zeta^{-1} \psi), 0, 1)} - \frac{1}{2} \right),
\]

and thus

\[
P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{s.t. } X_{\gamma_1, t} \neq X_{\gamma_2, t} \right)
\]

\[
= \int_0^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^2 \pi^2}{2\zeta^2} \right) \right) \frac{\zeta^{-2} \psi}{\ln 2} \left( 2^{-\min(\log_2(\zeta^{-1} \psi), 0, 1)} - \frac{1}{2} \right) d\zeta
\]

\[
= \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^2 \pi^2}{2\zeta^2} \right) \right) \frac{\zeta^{-2} \psi}{\ln 2} \left( \min \left\{ \zeta^{-1} \psi^{-1}, 1 \right\} - \frac{1}{2} \right) d\zeta.
\]
We then prove Corollary 2.3.

**Proof of Corollary 2.3.** We first prove Corollary 2.3.1. By Theorem 2.2,

\[
P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
= \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^2 \pi^2}{2 \zeta^2} \right) \right) \frac{\zeta^{-2}}{\ln 2} \left( \min \{ \zeta \psi^{-1}, 1 \} - \frac{1}{2} \right) d\zeta \\
= \psi \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} \left( \exp \left( -\frac{(2k-1)^2 \pi^2}{2 \zeta^2} \right) - \exp \left( -\frac{(2k)^2 \pi^2}{2 \zeta^2} \right) \right) \right) \frac{\zeta^{-2}}{2 \ln 2} d\zeta \\
\leq \psi \int_{0}^{\infty} \left( \sum_{k=1}^{\infty} \left( \exp \left( -\frac{(2k-1)^2 \pi^2}{2 \zeta^2} \right) - \exp \left( -\frac{(2k)^2 \pi^2}{2 \zeta^2} \right) \right) \right) \frac{\zeta^{-2}}{2 \ln 2} d\zeta \\
= \psi \int_{0}^{\infty} \left( \sum_{k=1}^{\infty} \left( \frac{1}{\sqrt{2\pi}(2k-1)} - \frac{1}{\sqrt{2\pi}(2k)} \right) \right) d\zeta \\
= \frac{\psi}{\sqrt{2\pi}},
\]

where (a) is by Fubini’s theorem, and (b) is by substituting \( \tau = 1/\zeta \).

For Corollary 2.3.3, if \( \psi \geq 2 \), by (11),

\[
P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
= P \left( \sup_{t \leq s} Y_t \leq \sqrt{s} Z \right) \\
= P \left( \sup_{t \leq 1} Y_t \leq Z \right) \\
\leq P \left( \sup_{t \leq 1} B_{1,t} \leq Z \text{ and } \sup_{t \leq 1} B_{2,t} \leq Z \text{ and } \sup_{t \leq 1} B_{3,t} \leq Z \right) \\
= \mathbb{E} \left[ \left( \text{erf} \left( \frac{Z}{\sqrt{2}} \right) \right)^3 \right] \\
\leq \mathbb{E} \left[ \text{erf} \left( \frac{\exp(\ln Z)}{\sqrt{2}} \right) \right] \\
\leq \text{erf} \left( \frac{\exp(\mathbb{E} [\ln Z])}{\sqrt{2}} \right),
\]

where in (a), we let \( Y_t \overset{d}{=} \sqrt{B_{1,t}^2 + B_{2,t}^2 + B_{3,t}^2} \), where \( \{B_{1,t}\}, \{B_{2,t}\}, \{B_{3,t}\} \) are independent Brownian motions, and (b) is because \( \gamma \mapsto \text{erf}(\exp(\gamma)/\sqrt{2}) \) is concave for \( \gamma \geq 0 \) (note that \( Z \geq \psi/2 \geq 1 \)). We have

\[
\mathbb{E} [\ln Z] \\
= \int_{\psi/2}^{\infty} (\ln \zeta) \frac{\zeta^{-2}}{\ln 2} \left( \min \{ \zeta \psi^{-1}, 1 \} - \frac{1}{2} \right) d\zeta \\
= \frac{\psi \ln \psi + \psi + \psi (\ln \psi)^2}{2 \psi \ln 2} - \frac{\psi \ln(\psi/2) + \psi + (\psi/2)(\ln(\psi/2))^2}{2(\psi/2) \ln 2} + \frac{\psi}{2 \ln 2} \cdot \frac{\ln \psi + 1}{\psi}
\]
\[
\begin{align*}
&= \frac{\ln \psi + 1 + (\ln \psi)^2}{2 \ln 2} - \frac{2 \ln \psi - 2 \ln 2 + 2 + (\ln \psi - \ln 2)^2}{2 \ln 2} + \frac{\ln \psi + 1}{2 \ln 2} \\
&= \frac{1}{2 \ln 2} ((\ln \psi)^2 - (\ln \psi - \ln 2)^2) + 1 \\
&= \frac{1}{2 \ln 2} (2 \ln \psi - \ln 2) \ln 2 + 1 \\
&= \ln \psi - \frac{\ln 2}{2} + 1.
\end{align*}
\]

Therefore,
\[
P\left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s, X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
\leq \text{erf} \left( \frac{1}{\sqrt{2}} \exp \left( \ln \psi - \frac{\ln 2}{2} + 1 \right) \right) \\
= \text{erf} \left( \frac{e\psi}{2} \right).
\]

If \( \psi < 2 \), then
\[
P\left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s, X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
\leq \frac{\psi}{\sqrt{2\pi}} \\
\leq \text{erf} \left( \frac{e\psi}{2} \right).
\]

The result follows.

For Corollary 2.3.2, for any \( 1 < \delta \leq 2 \),
\[
P\left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s, X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
\leq E \left[ \left( \text{erf} \left( \frac{Z}{\sqrt{2}} \right) \right)^3 \right] \\
\leq 1 - \left( 1 - \left( \text{erf} \left( \frac{\psi \delta}{2\sqrt{2}} \right) \right)^3 \right) P(Z \leq \psi \delta / 2) \\
= 1 - \left( 1 - \left( \text{erf} \left( \frac{\psi \delta}{2\sqrt{2}} \right) \right)^3 \right) \int_{\psi/2}^{\psi \delta \sqrt{2}} \frac{\zeta - \psi}{\ln 2} \left( \min \{ \zeta \psi^{-1}, 1 \} - \frac{1}{2} \right) \frac{d\zeta}{\ln 2} \\
= 1 - \left( 1 - \left( \text{erf} \left( \frac{\psi \delta}{2\sqrt{2}} \right) \right)^3 \right) \frac{\ln \delta + \delta^{-1} - 1}{\ln 2}
\]

If \( \psi \geq 2\sqrt{2} \), substituting \( \delta = 1 + 8/\psi^2 \), we have
\[
P\left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s, X_{\gamma_1, t} \neq X_{\gamma_2, t} \right) \\
\leq 1 - \left( 1 - \left( \text{erf} \left( \frac{\psi + 8/\psi}{2\sqrt{2}} \right) \right)^3 \right) \frac{\ln (1 + 8/\psi^2) + (1 + 8/\psi^2)^{-1} - 1}{\ln 2} \\
\leq 1 - \left( 1 - \text{erf} \left( \frac{\psi + 8/\psi}{2\sqrt{2}} \right) \right) \frac{\ln (1 + 8/\psi^2) + (1 + 8/\psi^2)^{-1} - 1}{\ln 2}.
\]
Note that \( \ln(1 + 8/\psi^2) + (1 + 8/\psi^2)^{-1} - 1 = \Omega(\psi^{-4}) \) as \( \psi \to \infty \), whereas \( 1 - \text{erf}(x) \to 0 \) exponentially as \( x \to \infty \). Therefore there exists a function \( \kappa : \mathbb{R}^+ \to \mathbb{R}^+ \) such that \( \lim_{t \to \infty} \kappa(t) = 0 \), and

\[
P \left( \exists \gamma_1, \gamma_2 \in [\alpha, \beta], t \geq s \text{ s.t. } X_{\gamma_1, t} \neq X_{\gamma_2, t} \right)
\leq \text{erf} \left( \frac{(1 + \kappa(t))\psi}{2\sqrt{2}} \right).
\]

This implies that

\[
\lim_{p \to 0} F_{\alpha, \beta}^{-1}(p) = 1.
\]

3. Nonexistence of a Pairwise Maximal Coupling

In this section, we show that there does not exist a pairwise maximal coupling \( \{\tilde{X}_{\alpha, t}\}_{t \geq 0} \) of \( \{\text{BM}(\alpha)\}_{\alpha \in \mathbb{R}} \), that is, one in which every pair \( \{\tilde{X}_{\alpha, t}\}_{t \geq 0}, \{\tilde{X}_{\beta, t}\}_{t \geq 0} \) is a maximal coupling, i.e.,

\[
P \left( \exists t \geq s \text{ s.t. } \tilde{X}_{\alpha, t} \neq \tilde{X}_{\beta, t} \right) = \text{erf} \left( \frac{|\alpha - \beta|}{2\sqrt{2}s} \right).
\]

(14)

Note that both (14) and the expression in Theorem 2.2 depend only on \( \psi := |\alpha - \beta|/\sqrt{s} \).

**Definition 3.1.** We say that a function \( \hat{h} : [0, \infty) \to [0, 1] \) is an attainable failure probability bound if there exists a coupling \( \{\tilde{X}_{\alpha, t}\}_{t \geq 0} \) of \( \{\text{BM}(\alpha)\}_{\alpha \in \mathbb{R}} \) such that

\[
P \left( \exists t \geq s \text{ s.t. } \tilde{X}_{\alpha, t} \neq \tilde{X}_{\beta, t} \right) \leq \hat{h} \left( \frac{|\alpha - \beta|}{\sqrt{s}} \right)
\]

for all \( \alpha, \beta, s \in \mathbb{R}, s > 0 \). We say that \( c > 0 \) is an attainable multiplicative gap if \( x \mapsto \text{erf}(x\sqrt{c}/(2\sqrt{2})) \) is an attainable failure probability bound.

One attainable failure probability bound is given in Theorem 2.2. Corollary 2.3.3 implies that \( 2e^2 \) is an attainable multiplicative gap. A pairwise maximal coupling exists if and only if \( x \mapsto \text{erf}(x/(2\sqrt{2})) \) is an attainable failure probability bound, or equivalently, 1 is an attainable multiplicative gap.

We now prove a lower bound on any attainable failure probability bound which implies a lower bound on the attainable multiplicative gap. This implies the nonexistence of a pairwise maximal coupling.

**Theorem 3.2.** If \( h \) is a failure probability bound, then for any \( 0 < s < t \), we have

\[
\tilde{h} \left( \frac{2}{\sqrt{t}} \right) + \hat{h} \left( \frac{2}{\sqrt{s}} \right) + 2\hat{h} \left( \frac{1}{\sqrt{s}} \right)
\geq \text{erf} \left( \frac{1}{\sqrt{2t}} \right) - \text{erf} \left( \frac{1}{\sqrt{2s}} \right) - \text{erf} \left( \frac{1}{4\sqrt{2(t-s)}} \right).
\]

where \( \tilde{h}(x) := h(x) - \text{erf}(x/(2\sqrt{2})) \), and \( \text{erfc}(x) := 1 - \text{erf}(x) \). Hence \( x \mapsto \text{erf}(x/(2\sqrt{2})) \) is not an attainable failure probability bound. Moreover, 1.0025 is not an attainable multiplicative gap.

**Proof.** For any \( \alpha < \beta, s > 0 \), we have

\[
P(X_{\beta, s} \leq (\alpha + \beta)/2 \text{ and } X_{\beta, s} \neq X_{\alpha, s}) + P(X_{\alpha, s} \geq (\alpha + \beta)/2 \text{ and } X_{\beta, s} \neq X_{\alpha, s})
\]

\[
= P(X_{\beta, s} \leq (\alpha + \beta)/2) + P(X_{\alpha, s} \geq (\alpha + \beta)/2)
\]

\[
- (P(X_{\beta, s} \leq (\alpha + \beta)/2 \text{ and } X_{\beta, s} = X_{\alpha, s}) + P(X_{\alpha, s} \geq (\alpha + \beta)/2 \text{ and } X_{\beta, s} = X_{\alpha, s}))
\]
\[ \begin{align*}
& \leq P(X_{\beta,s} \leq (\alpha + \beta)/2) + P(X_{\alpha,s} \geq (\alpha + \beta)/2) - P(X_{\beta,s} = X_{\alpha,s}) \\
& \overset{(a)}{\leq} 1 - \text{erf} \left( \frac{\alpha - \beta}{2\sqrt{2}s} \right) - \left( 1 - h \left( \frac{\alpha - \beta}{\sqrt{s}} \right) \right) \\
& = \tilde{h} \left( \frac{\alpha - \beta}{\sqrt{s}} \right),
\end{align*} \]

where (a) is because \( X_{\alpha,s} \sim N(\alpha, s) \), \( X_{\beta,s} \sim N(\beta, s) \) and by the definition of the failure probability bound.

Let \( 0 < s < t \). We have

\[ \begin{align*}
P(X_{1,s} - X_{-1,s} \leq 1/2 \text{ and } X_{1,s} \neq X_{-1,s}) \\
& \leq P(X_{1,s} \leq 0 \text{ and } X_{1,s} \neq X_{-1,s}) + P(X_{-1,s} \geq 0 \text{ and } X_{1,s} \neq X_{-1,s}) \\
& + P(X_{1,s} \leq 1/2 \text{ and } X_{-1,s} \geq -1/2 \text{ and } X_{1,s} \neq X_{-1,s}) \\
& \overset{(a)}{\leq} P(X_{1,s} \leq 0 \text{ and } X_{1,s} \neq X_{-1,s}) + P(X_{-1,s} \geq 0 \text{ and } X_{1,s} \neq X_{-1,s}) \\
& + P(X_{1,s} \leq 1/2 \text{ and } X_{1,s} \neq X_{0,s}) + P(X_{-1,s} \geq -1/2 \text{ and } X_{-1,s} \neq X_{0,s}) \\
& \overset{(b)}{\leq} \tilde{h}(2/\sqrt{s}) + 2\tilde{h}(1/\sqrt{s}),
\end{align*} \]

where (a) is because if \( X_{1,s} \neq X_{-1,s} \), then either \( X_{1,s} \neq X_{0,s} \) or \( X_{-1,s} \neq X_{0,s} \), and (b) is by applying (15) on \( (\alpha, \beta, s) \leftarrow (-1, 1, s), (0, 1, s) \) and \( (-1, 0, s) \) respectively. Hence,

\[ \begin{align*}
P(X_{1,t} = X_{-1,t} \text{ and } X_{1,s} \neq X_{-1,s}) \\
& \leq P(X_{1,t} - X_{-1,t} \leq 1/2 \text{ and } X_{1,s} \neq X_{-1,s}) \\
& + P(X_{1,t} - X_{1,s} \leq 1/4 \text{ and } X_{1,t} - X_{-1,s} \geq 1/4) \\
& \leq \tilde{h}(2/\sqrt{s}) + 2\tilde{h}(1/\sqrt{s}) + \text{erfc} \left( \frac{1/4}{\sqrt{2(t-s)}} \right).
\end{align*} \]

Therefore,

\[ \begin{align*}
1 - \tilde{h}(2/\sqrt{t}) \\
& \leq P(X_{1,t} = X_{-1,t}) \\
& \leq P(X_{1,s} = X_{-1,s}) + P(X_{1,t} = X_{-1,t} \text{ and } X_{1,s} \neq X_{-1,s}) \\
& \leq \text{erfc} \left( \frac{2}{2\sqrt{2}s} \right) + \tilde{h}(2/\sqrt{s}) + 2\tilde{h}(1/\sqrt{s}) + \text{erfc} \left( \frac{1/4}{\sqrt{2(t-s)}} \right) \\
& = \text{erfc} \left( \frac{1}{\sqrt{2}s} \right) + \text{erfc} \left( \frac{1}{4\sqrt{2(t-s)}} \right) + \tilde{h}(2/\sqrt{s}) + 2\tilde{h}(1/\sqrt{s}).
\end{align*} \]

Hence,

\[ \begin{align*}
\tilde{h} \left( \frac{2}{\sqrt{t}} \right) + \tilde{h} \left( \frac{2}{\sqrt{s}} \right) + 2\tilde{h} \left( \frac{1}{\sqrt{s}} \right) \\
& \geq \text{erfc} \left( \frac{1}{\sqrt{2t}} \right) - \text{erfc} \left( \frac{1}{\sqrt{2s}} \right) - \text{erfc} \left( \frac{1}{4\sqrt{2(t-s)}} \right).
\end{align*} \]

Note that the above lower bound can be positive (e.g. it is at least 0.0019 when \( s = 0.33, t = 0.3348 \)), and thus \( x \mapsto \text{erf}(x/2\sqrt{2}) \) is not an attainable failure probability bound.

It can be verified numerically that \( x \mapsto \text{erf}(x/\sqrt{c}/(2\sqrt{2})) \) does not satisfy the above inequality when \( c = 1.0025, s = 0.2361, t = 0.2408 \). Hence 1.0025 is not an attainable multiplicative gap.
We conjecture that the dyadic grand coupling is optimal in the following sense.

Conjecture 3.3. If $h$ is a failure probability bound, then for any $\psi > 0$, we have

$$h(\psi) \geq \int_{\psi/2}^{\infty} \left( \sum_{k=1}^{\infty} 2(-1)^{k+1} \exp \left( -\frac{k^2\pi^2}{2\zeta^2} \right) \right)^{1/2} \frac{\zeta^{-\psi}}{\ln 2} \left( \min \{ \zeta^{-1}, 1 \} - \frac{1}{2} \right) d\zeta,$$

i.e., the attainable failure probability bound given in Theorem 2.2 is pointwise optimal.

Loosely speaking, the dyadic grand coupling is “locally a reflection coupling”, in the sense that the coupled processes after the time of each coalescence point can be obtained by performing the reflection coupling between adjacent pairs of coalescence points (see Figure 2). In Conjecture 3.3, we raise the question whether such “locally optimal” coupling is globally optimal.

It may also be of interest to find the smallest attainable multiplicative gap. Theorem 3.2 and the numerical evidence in Figure 4 show that the infimum of the set of attainable multiplicative gaps is between 1.0025 and 1.5.

4. Conclusions and Discussion

We constructed a coupling of $\{BM(\alpha)\}_{\alpha \in \mathbb{R}}$, such that the coupling for any pair of the coupled processes is close to being maximal. While it is shown that a pairwise exactly maximal coupling does not exist, we conjecture that our coupling is optimal among couplings of $\{BM(\alpha)\}_{\alpha \in \mathbb{R}}$ in the sense of attainable failure probability bounds.

One future direction is to generalize the construction to Brownian motions in $\mathbb{R}^n$. While we can couple each coordinate independently using the dyadic grand coupling, this may not be the optimal construction.

Another direction is to consider Brownian motions with initial distributions (rather than fixed starting points), i.e., the collection of processes is $\{BM(P)\}_{P \in \mathcal{P}(\mathbb{R})}$, where $\mathcal{P}(\mathbb{R})$ is the set of distributions over $\mathbb{R}$, and $BM(P)$ is the Brownian motion with initial distribution $P$. One simple construction is to first couple the starting point by the quantile coupling, then apply the dyadic grand coupling, i.e., $X_{P,0} := F^{-1}_P(U)$, where $U \sim \text{Unif}[0,1]$, and $X_{P,t} := X_{X_{P,0},t}$ for $t > 0$, where $X_{X_{P,0},t}$ is given by the dyadic grand coupling with starting point $X_{P,0}$. Another construction is to use the sequential Poisson functional representation [22] instead of the quantile coupling, since it is more suitable for minimizing concave costs (it is shown in Appendix B that the probability of failure in Theorem 2.2 is concave in $|\alpha - \beta|$).

Appendix A: Proof of the Claim in Definition 2.1

We first prove that for any $\theta \in [0,1]$, $\alpha \in \mathbb{R}$, we have $G_{\alpha,j} = W_j$ for all sufficiently large $j$, as long as $\sup \{ j : W_j = 1 \} = \sup \{ j : W_j = -1 \} = \infty$. Let $k_1 \in \mathbb{Z}$ be such that $2^{k_1} > 4|\alpha|$ and $W_{k_1} = 1$, and $k_{-1} \in \mathbb{Z}$ be such that $2^{k_{-1}} > 4|\alpha|$ and $W_{k_{-1}} = -1$. Assume $j \geq \max \{ k_1, k_{-1} \}$. We have

$$\alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} \leq \alpha + \sum_{k \leq j-1, k \neq k_1} 2^{k+\theta-1} - 2^{k_1+\theta-1} = \alpha + 2^{j+\theta-1} - 2^{k_1+\theta} < 2^{j+\theta-1},$$

where the last inequality is by the definition of $k_1$. Similarly, $\alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} > -2^{j+\theta-1}$. Hence,

$$\left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^{j+\theta-1} \right) \mod 2^{j+\theta+1} \in [0, 2^{j+\theta}).$$
and \(G_{\theta,\alpha,j} = W_j\).

We then prove (7). We will prove by induction that for all \(j \in \mathbb{Z}\),

\[
2^{-(j+\theta)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 \right)
\]

\[
= 2^{-(j+\theta)} \sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k}) 2^{k+\theta-1}.
\]

If \(j \geq \max\{k_1, k_{-1}\}\), then \(-2^j + \theta - 1 < \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} < 2^j + \theta - 1\), and also \(W_k = G_{\theta,\alpha,k}\) for \(k \geq j\), and thus both sides in the induction hypothesis are 0.

Assume the induction hypothesis is true for \(j + 1\). If \(W_j = G_{\theta,\alpha,j}\), then

\[
\left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 \right) \mod 2^j + \theta + 1 \in [0, 2^j + \theta),
\]

and hence

\[
2^{-(j+\theta)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 \right)
\]

\[
\equiv_2 \left[ 2^{-(j+\theta+1)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 + (1 - W_j) 2^j + \theta - 1 \right) \right]
\]

\[
= 2 \left[ 2^{-(j+\theta+1)} \left( \alpha - \sum_{k=-\infty}^{j} W_k 2^{k+\theta-1} + 2^j + \theta \right) \right]
\]

\[
\equiv_2 2 \cdot 2^{-(j+1+\theta)} \sum_{k=j+1}^{\infty} (W_k - G_{\theta,\alpha,k}) 2^{k+\theta-1}
\]

\[
= 2^{-(j+\theta)} \sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k}) 2^{k+\theta-1},
\]

where (a) is because \((1 - W_j) 2^j + \theta - 1 \in [0, 2^j + \theta)\), and (b) is by the induction hypothesis for \(j + 1\). Therefore the induction hypothesis holds for \(j\). If \(W_j = -G_{\theta,\alpha,j}\), then

\[
\left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 \right) \mod 2^j + \theta + 1 \in [2^j + \theta, 2^j + \theta + 1),
\]

and hence

\[
2^{-(j+\theta)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 \right)
\]

\[
\equiv_2 2 \left[ 2^{-(j+\theta+1)} \left( \alpha - \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} + 2^j + \theta - 1 + (1 - W_j) 2^j + \theta - 1 \right) \right] + W_j
\]

\[
= 2 \left[ 2^{-(j+\theta+1)} \left( \alpha - \sum_{k=-\infty}^{j} W_k 2^{k+\theta-1} + 2^j + \theta \right) \right] + W_j
\]

\[
= 2 \cdot 2^{-(j+1+\theta)} \sum_{k=j+1}^{\infty} (W_k - G_{\theta,\alpha,k}) 2^{k+\theta-1} + W_j
\]
\[
2^{-(j+\theta)} \sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k})2^{k+\theta-1},
\]

where (a) can be deduced by considering whether \( W_j = 1 \) or \(-1 \). Therefore the induction hypothesis holds for \( j \).

Hence the induction hypothesis holds for all \( j \in \mathbb{Z} \), and

\[
\sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k})2^{k+\theta-1} - 2^{j+\theta-1} \leq \alpha \leq \sum_{k=-\infty}^{j-1} W_k 2^{k+\theta-1} < \sum_{k=j}^{\infty} (W_k - G_{\theta,\alpha,k})2^{k+\theta-1} + 2^{j+\theta-1}.
\]

Letting \( j \to -\infty \), we have \( \sum_{j=-\infty}^{\infty} (W_j - G_{\theta,\alpha,j})2^{j+\theta-1} = \alpha \).

**Appendix B: Proof that the Expression in Theorem 2.2 is concave in \( |\alpha - \beta| \)**

Let \( h(\psi) := P(\exists \gamma_1, \gamma_2 \in [0, \psi], t > 1 \text{ s.t. } X_{\gamma_1,t} \neq X_{\gamma_2,t}) \) be as given in Theorem 2.2. Let \( 0 < \psi_1 < \psi_2 \). Fix \( l > \psi_2 \). Let \( (A_1, B_1), (A_2, B_2), \ldots, (A_N, B_N) \subseteq [0, l] \) be maximal open intervals with length at least \( \psi_1 \), sorted in ascending order, such that \( \{X_{\gamma,t}, t \geq 1\} \) is constant within each interval (i.e., for any \( i = 1, \ldots, N \), we have \( B_i - A_i \geq \psi_1 \), \( \{X_{\gamma_1,t}, t \geq 1\} = \{X_{\gamma_2,t}, t \geq 1\} \) for any \( \gamma_1, \gamma_2 \in (A_i, B_i) \), and any open interval in \([0, l] \) that is a proper superset of \((A_i, B_i)\) does not have this property). For any \( \rho \in [0, 1] \), let \( \psi := \rho \psi_1 + (1 - \rho) \psi_2 \). We have

\[
h(\psi) = \frac{1}{l - \psi} \int_0^{l-\psi} P(\exists \gamma_1, \gamma_2 \in [x, x + \psi], t > 1 \text{ s.t. } X_{\gamma_1,t} \neq X_{\gamma_2,t}) \, dx
\]

\[
= 1 - \frac{1}{l - \psi} \mathbb{E} \left[ \sum_{i=1}^{N} \max\{B_i - A_i - \psi, 0\} \right]
\]

\[
\geq 1 - \frac{1}{l - \psi} \mathbb{E} \left[ \sum_{i=1}^{N} (\rho \max\{B_i - A_i - \psi_1, 0\} + (1 - \rho) \max\{B_i - A_i - \psi_2, 0\}) \right]
\]

\[
= 1 - \frac{1}{l - \psi} \left( \rho \mathbb{E} \left[ \sum_{i=1}^{N} \max\{B_i - A_i - \psi_1, 0\} \right] + (1 - \rho) \mathbb{E} \left[ \sum_{i=1}^{N} \max\{B_i - A_i - \psi_2, 0\} \right] \right)
\]

\[
= 1 - \frac{1}{l - \psi} \left( \rho (l - \psi_1)(1 - h(\psi_1)) + (1 - \rho)(l - \psi_2)(1 - h(\psi_2)) \right),
\]

where (a) is by the convexity of \( x \mapsto \max\{\gamma - x, 0\} \). Letting \( l \to \infty \), we have \( h(\psi) \geq \rho h(\psi_1) + (1 - \rho) h(\psi_2) \). Hence \( h \) is concave on \((0, \infty)\). Since \( h \) is non-decreasing, \( h \) is concave on \([0, \infty)\).

**Acknowledgements**

The authors acknowledge support from the NSF grants CNS-1527846, CCF-1618145, CCF-1901004, the NSF Science & Technology Center grant CCF-0939370 (Science of Information), and the William and Flora Hewlett Foundation supported Center for Long Term Cybersecurity at Berkeley. The work of C. T. Li is partially supported by the Direct Grant for Research at the Chinese University of Hong Kong.
References