

# OPTIMAL EMBEDDINGS BY UNBIASED SHIFTS OF BROWNIAN MOTION

PETER MÖRTERS AND ISTVÁN REDL

ABSTRACT. An unbiased shift of the two-sided Brownian motion  $(B_t: t \in \mathbb{R})$  is a random time  $T$  such that  $(B_{T+t}: t \in \mathbb{R})$  is still a two-sided Brownian motion. Given a pair  $\mu, \nu$  of orthogonal probability measures, an unbiased shift  $T$  solves the embedding problem, if  $B_0 \sim \mu$  implies  $B_T \sim \nu$ . A solution to this problem was given by Last et al. (2014), based on earlier work of Bertoin and Le Jan (1992), and Holroyd and Liggett (2001). In this note we show that this solution minimises  $\mathbb{E}\psi(T)$  over all nonnegative unbiased solutions  $T$ , simultaneously for all nonnegative, concave functions  $\psi$ . Our proof is based on a discrete concavity inequality that may be of independent interest.

**MSC (2010):** 60J65 (primary) 39B62, 49J55, 60G40 (secondary).

**Keywords:** Brownian motion, Skorokhod embedding, shift-coupling, unbiased shift, optimal transport, concavity.

**1. Introduction.** A random time  $T$  is called an *unbiased shift* of the two sided Brownian motion  $(B_t: t \in \mathbb{R})$  if the shifted process  $(B_{T+t}: t \in \mathbb{R})$  is again a two-sided Brownian motion. This notion was introduced and studied by Last et al. in [10], where the additional requirement of measurability of  $T$  with respect to the process  $(B_t: t \in \mathbb{R})$  is made, which we drop here for greater generality. The concept of unbiased shifts goes back to a similar idea for coin tosses, known as the extra head problem in [7, 12, 9], and, more generally, the concept of shift-couplings, see the work of Thorisson [15].

In [10] the authors solve the *embedding problem* for unbiased shifts. Namely, given two orthogonal probability measures  $\mu, \nu$  on the real line such that  $B_0 \sim \mu$  they construct a nonnegative unbiased shift  $T_*$  such that  $B_{T_*} \sim \nu$ . The solution of [10] can easily be described explicitly. Let  $(L_t^x: x \in \mathbb{R}, t \geq 0)$  be a continuous version of the local time for  $(B_t: t \geq 0)$ . We use this to build two continuous additive functionals by letting

$$L_t^\nu := \int L_t^x \nu(dx), \text{ and } L_t^\mu := \int L_t^x \mu(dx),$$

and obtain the solution as

$$T_* := \inf\{t > 0: L_t^\nu = L_t^\mu\}. \tag{1.1}$$

Note that  $T_*$  occurred in the context of one-sided (Skorokhod) embedding problems in the work of Bertoin and Le Jan [2], is reminiscent of extra head schemes in [12, 9] or [14], and turns out to be closely related to allocation and transport problems, see [8, 1].

The present paper is concerned with the problem whether this natural and explicit solution of the embedding problem is optimal in the sense that it minimises certain moments. To see which moments should be looked at, Last et al. [10] have investigated finiteness of moments for unbiased shifts, and have shown that, for any unbiased shift  $T$  embedding a measure orthogonal to the initial distribution, the random variable  $|T|$  has infinite square-root moment. Under the additional assumption that  $T$  is a nonnegative stopping time they obtain that  $T$  has an infinite fourth-root moment. It is conjectured in [10] that this holds without the assumption that  $T$  is a stopping time, and this conjecture will be confirmed in this paper, see Remark 1(c) below. The results of [10] show that in order to understand optimality we need to focus on moments of fractional order smaller than  $\frac{1}{2}$ , or more generally moments taken with respect to concave gauge functions.

We denote by  $\mathbb{P}_\mu, \mathbb{E}_\mu$  probability and expectation on a probability space supporting a two-sided Brownian motion  $(B_t: t \in \mathbb{R})$  with  $B_0 \sim \mu$ . We can now state our main result.

**Theorem 1.** *Assume that  $\mu$  and  $\nu$  are two orthogonal probability measures on  $\mathbb{R}$ . For any non-negative unbiased shift  $T$  embedding  $\nu$  and all  $\psi: [0, \infty) \rightarrow [0, \infty)$  concave,*

$$\mathbb{E}_\mu \psi(T_*) \leq \mathbb{E}_\mu \psi(T),$$

where  $T_*$  is the unbiased shift constructed in (1.1).

*Remark 1*

- (a) This result is closely related to a similar optimality result in [14] for the case of discrete-time Markov chains. Although our proof relies on a discrete approximation, we have been unable to derive Theorem 1 directly from the results of [14]. Instead, we use a concavity inequality, stated as Lemma 3.2(b) below, that may be of independent interest.
- (b) We make no assumptions on measurability of  $T$  with respect to the Brownian motion, i.e.,  $T$  is allowed to use additional randomisation beyond that taken from the Brownian motion.
- (c) As  $T_*$  is a stopping time, we have  $\mathbb{E}_\mu T_*^{1/4} = \infty$  by Theorem 8.1 in [10], and hence

$$\mathbb{E}_\mu T^{1/4} = \infty$$

for all non-negative solutions  $T$  of the embedding problem.

- (d) Under mild conditions on  $\mu, \nu$  we have  $\mathbb{E}_\mu T_*^\alpha < \infty$  for  $\alpha < \frac{1}{4}$  by Theorem 8.2 in [10].
- (e) The result is strongly reminiscent of optimality results for Skorokhod embeddings, as given, for example, in the classical papers [3, 4]. The use of optimal transport ideas in connection to optimal Skorokhod embeddings is also the topic of a lot of current research, of which the paper [1] is a major highlight.

*Remark 2* If  $\mu$  and  $\nu$  are not orthogonal we can write  $\mu = \tilde{\mu} + \mu \wedge \nu$  and  $\nu = \tilde{\nu} + \mu \wedge \nu$  for orthogonal sub-probability measures  $\tilde{\mu}, \tilde{\nu}$ . Denoting the total mass of both measures by  $\rho$  one can enlarge the probability space to include a Bernoulli variable  $U$ , depending only on  $B_0$ , with  $\mathbb{P}_\mu(U = 1) = \rho$  such that on the event  $U = 1$  we have  $B_0 \sim \frac{1}{\rho} \tilde{\mu}$  and on the event  $U = 0$  we have  $B_0 \sim \frac{1}{1-\rho} (\mu \wedge \nu)$ . Then

$$T_* = U \inf\{t > 0: L_t^{\tilde{\nu}} = L_t^{\tilde{\mu}}\}$$

is a non-negative unbiased shift embedding  $\nu$ , which satisfies the inequality in Theorem 1. This follows by splitting the expectations according to the value of  $U$ , and using monotonicity of  $\psi$  on the event  $U = 0$ , and applying the theorem to the orthogonal probability measures  $\frac{1}{\rho} \tilde{\mu}, \frac{1}{\rho} \tilde{\nu}$  on the event  $U = 1$ .

**2. Preliminaries and outline of the proof.** We recall the framework of [10], which will be used as a main reference throughout this paper. Let  $(\Omega, \mathcal{A}, \mathbb{P}_0)$  be such that  $\Omega$  is the space of continuous functions  $\omega: \mathbb{R} \rightarrow \mathbb{R}$ , equipped with the Borel  $\sigma$ -algebra  $\mathcal{A}$  and the distribution  $\mathbb{P}_0$  of two-sided Brownian motion with  $\omega_0 = 0$ . For all  $x \in \mathbb{R}$  we define  $\mathbb{P}_x$  to be the distribution of  $\omega + x$  and introduce the  $\sigma$ -finite measure  $\mathbb{P} := \int \mathbb{P}_x dx$ , which is invariant under the shifts  $\omega \mapsto s\omega$  defined by  $(s\omega)(\cdot) = \omega(\cdot - s)$ , for all  $s \in \mathbb{R}$ . As usual, expectations with respect to  $\mathbb{P}, \mathbb{P}_x$  will be denoted by  $\mathbb{E}, \mathbb{E}_x$ , respectively.

Let  $(\ell^x: x \in \mathbb{R})$  be a continuous version of the family of local times of  $\omega$  at level  $x$ , and for any locally finite measure  $\nu$  on  $\mathbb{R}$  set  $\ell^\nu := \int \ell^x \nu(dx)$ . We note that the local times  $\ell^x$  and  $\ell^\nu$  are random measures and depend on  $\omega \in \Omega$ . When necessary, this dependence will be indicated as  $\ell_\omega$ . We may and shall assume that  $\ell_{z+\omega}^x = \ell_\omega^{x-z}$ , for all  $x, z \in \mathbb{R}$ , for  $\mathbb{P}$ -almost every  $\omega$ .

A *transport rule* is a Markov kernel  $\theta: \Omega \times \mathbb{R} \times \mathcal{B} \rightarrow [0, 1]$ , where  $\mathcal{B}$  is the Borel  $\sigma$ -algebra on  $\mathbb{R}$ . We use the notation  $\theta_\omega(ds, dt) = \theta_\omega(s, dt)\ell^\mu(ds)$ , and  $\theta_\omega(A, B) = \int_A \theta_\omega(s, B)\ell^\mu(ds)$  for any  $A, B \in \mathcal{B}$ . An interpretation of the Markov kernels  $\theta(s, \cdot)$  is that each site  $s \in \mathbb{R}$  sends out unit mass to the real line. A transport rule  $\theta$  is called *translation invariant* if  $\theta_\omega(s, A) = \theta_{t\omega}(s-t, A-t)$ , for any  $(s, A) \in \mathbb{R} \times \mathcal{B}$ ,  $t \in \mathbb{R}$ , for  $\mathbb{P}$ -almost every  $\omega$ .

We now briefly summarise results of [10], which are derived from the abstract results of [11]. Given a translation invariant transport rule  $\theta$  we obtain an unbiased shift  $T$  by letting

$$\mathbb{P}_\mu(\omega \in A, T \in B) = \iint_A \theta_\omega(0, B) \mathbb{P}_x(d\omega) \mu(dx), \text{ for } A \in \mathcal{A}, B \in \mathcal{B}. \quad (2.1)$$

Conversely, given an unbiased shift  $T$  we can construct a transport rule  $\theta$  by letting  $\theta_\omega(s, B) = \mathbb{P}_\mu(T \in B - s \mid s\omega)$ , for  $s \in \mathbb{R}, B \in \mathcal{B}, \omega \in \Omega$ . We use a suitably regularised version of conditional probabilities so that  $\theta$  is a translation invariant transport rule and (2.1) holds.

The transport rules associated in this way with nonnegative unbiased shifts are *forward looking* in the sense that  $\theta_\omega(s, (-\infty, s)) = 0$  for all  $s \in \mathbb{R}$ . An unbiased shift  $T$  solves the embedding problem for a pair of orthogonal probability measures  $\mu, \nu$  if and only if the associated transport rule satisfies the *balancing property*

$$\int \theta_\omega(s, dt) \ell^\mu(ds) = \ell^\nu(dt) \quad \text{for } \mathbb{P}\text{-almost all } \omega.$$

If the unbiased shift is not randomised, i.e. a measurable function of  $\omega$ , then the associated  $\theta$  is an *allocation rule*, i.e. it is of the form  $\theta_\omega(s, A) = \mathbf{1}_A(\tau_\omega(s))$  for some  $\tau_\omega: \mathbb{R} \rightarrow \mathbb{R}$ . In this case each site  $s \in \mathbb{R}$  is assigned to a new site  $\tau(s) \in \mathbb{R}$ , where again we drop the dependence on  $\omega$  from the notation. The allocation rule associated to the random times  $T_*$  is given by

$$\tau_*(s) := \inf\{t > s: \ell^\mu[s, t] = \ell^\nu[s, t]\}, \quad \text{for all } s \in \mathbb{R}. \quad (2.2)$$

Let us now outline the proof of Theorem 1. The first part looks at what happens *pathwise*. With every transport rule we associate a local cost. Given a fixed  $\omega$ , we show that on carefully chosen intervals called excursions, the best possible cost is offered by the allocation rule  $\tau_*$ . We do this by proving an analogous discrete result and then taking a suitable limit. The second part of the proof uses *ergodic theory* to translate the local cost optimality of  $\tau_*$  into a result on the moments of  $T_*$ .

**3. Pathwise level.** In what follows, we fix a path  $\omega$ . An *excursion*  $\mathcal{E}$  is any interval of the form  $[a, \tau_*(a)]$ , for  $a \in \mathbb{R}$ . Observe that  $\tau_*$  maps  $\mathcal{E}$  onto itself. Our main goal is to show that the allocation rule  $\tau_*$  offers the best possible cost inside an arbitrary excursion  $\mathcal{E}$ . More precisely we will devote this section to proving the following proposition.

**Proposition 3.1.** *Given an excursion  $\mathcal{E}$ , for any forward-looking transport rule  $\theta$  balancing  $\ell^\mu$  and  $\ell^\nu$ ,*

$$\iint_{\mathcal{E} \times \mathbb{R}} \psi(t-s) \theta(ds, dt) + \iint_{\mathbb{R} \times \mathcal{E}} \psi(t-s) \theta(ds, dt) \geq 2 \int_{\mathcal{E}} \psi(\tau_*(s) - s) \ell^\mu(ds), \quad (3.1)$$

for all  $\psi: [0, \infty) \rightarrow [0, \infty)$  concave.

The left hand side in (3.1) is called the *cost* of the transport rule  $\theta$  over the interval  $\mathcal{E}$ . Note that  $\tau_*(\mathcal{E}) = \mathcal{E}$  and hence the right hand side is then the cost of the allocation rule given by  $\tau_*$  over the same interval. By subtracting a constant from both sides of the equation we may henceforth assume that  $\psi(0) = 0$ .

We start by establishing a similar result in a discrete setting. The inequality we establish below is of a general nature and may be of independent interest. A map  $\tau: A \rightarrow B$  between two discrete and

disjoint sets  $A, B \subset \mathbb{R}$  is the *stable allocation map* if

$$\tau(a) = \min \{b \in B : b > a, |B \cap [a, b]| = |A \cap [a, b]|\}.$$

**Lemma 3.2.** *Let  $a_1 > a_2 > a_3 > \dots$  and  $b_1 < b_2 < b_3 < \dots$  be disjoint real sequences, such that we have  $a_n \searrow -\infty$  and  $b_n \nearrow \infty$ .*

- (a) *The stable allocation map  $\tau: \{\dots, a_3, a_2, a_1\} \rightarrow \{b_1, b_2, b_3, \dots\}$  is well-defined and there exists  $N \in \mathbb{N}$  such that  $\tau(a_i) = b_i$  for all  $i \geq N$ .*  
 (b) *For every concave function  $\psi: [0, \infty) \rightarrow [0, \infty)$  and nonnegative matrix  $\pi = (\pi_{i,j}: i, j \in \mathbb{N})$  with the properties that*

- $\pi_{i,j} = 0$  if  $a_i > b_j$ ,
- $\sum_{j=1}^{\infty} \pi_{i,j} = 1$  for all  $i \in \{1, \dots, N\}$ ,
- $\sum_{i=1}^{\infty} \pi_{i,j} = 1$  for all  $j \in \{1, \dots, N\}$ ,

we have

$$\sum_{i=1}^N \sum_{j=1}^{\infty} \pi_{i,j} + \sum_{i=1}^{\infty} \sum_{j=1}^N \pi_{i,j} \psi(b_j - a_i) \geq 2 \sum_{i=1}^N \psi(\tau(a_i) - a_i). \quad (3.2)$$

Note that the sum on the left hand side of (3.2) is over pairs of positive integers  $i, j$  with  $i \wedge j \leq N$ , counting the contribution of pairs with  $i \vee j \leq N$  twice. In the special case that  $a_1 < b_1$  we have  $\tau(a_i) = b_i$  for all  $i$ , and hence  $N \in \mathbb{N}$  can be chosen arbitrarily. Then our result becomes the following general result, which to the best of our knowledge is new, too.

**Corollary 3.3.** *For every double-sided increasing sequence  $(a_n: n \in \mathbb{Z})$  that is unbounded from above and below, every stochastic matrix  $\pi = (\pi_{i,j}: i, j \in \mathbb{N})$ , and every concave function  $\psi: [0, \infty) \rightarrow [0, \infty)$  we have*

$$\sum_{i=1}^n \sum_{j=1}^{\infty} \pi_{i,j} + \sum_{i=1}^{\infty} \sum_{j=1}^n \pi_{i,j} \psi(a_i - a_{-j}) \geq 2 \sum_{i=1}^n \psi(a_i - a_{-i}).$$

*Proof of Lemma 3.2(a).* Given  $a \in A$  the set  $A \cap [a, \infty)$  is finite, but the set  $B \cap [a, \infty)$  is infinite. Hence, on the one hand, there exists  $b \in B$  such that  $|B \cap [a, b]| \geq |A \cap [a, \infty)| \geq |A \cap [a, b]|$ , while on the other hand  $|B \cap [a, a]| = 0 < 1 = |A \cap [a, a]|$ . This implies that there exists  $b' \in [a, b]$  with  $|B \cap [a, b']| = |A \cap [a, b']|$ , and hence  $\tau$  is well-defined.

We now show that there exists  $N \in \mathbb{N}$  such that, for all  $n \geq N$ , we have  $\tau(a_n) = b_n$ . If  $a_1 < b_1$  one can choose  $N = 1$ . Otherwise define an integer-valued function  $f: [b_1, \infty) \rightarrow \mathbb{R}$  by

$$f(x) = |A \cap [b_1, x]| - |B \cap [b_1, x]|.$$

Let  $M \in \mathbb{Z}$  be the minimum of  $f$  on  $[b_1, a_1]$ . Note that  $f(a_1) > M$  and on  $[a_1, \infty)$  the function  $f$  is decreasing to  $-\infty$  by downward jumps of size one. Hence there exists  $n > 1$  with  $f(b_n) = M - 1$ . Clearly, for all  $m \geq n$  we have  $|A \cap [a_m, b_m]| = m = |B \cap [a_m, b_m]|$  while, for  $j < m$ , we have  $|A \cap [a_m, b_j]| > j = |B \cap [a_m, b_j]|$ . Hence,  $\tau(a_m) = b_m$ , as required.  $\square$

*Proof of Lemma 3.2(b).* This is a variant of Lemma 5.3 in [14]. We say that four points  $a_k < a_i < b_j < b_l$  are crossed by  $\pi$  if  $\pi_{kj} > 0$  and  $\pi_{il} > 0$ . Such a crossing can be repaired by replacing the matrix  $\pi$  by a new matrix  $\pi'$  given by

$$\begin{aligned} \pi'_{kj} &= \pi_{kj} - (\pi_{kj} \wedge \pi_{il}), & \pi'_{il} &= \pi_{il} - (\pi_{kj} \wedge \pi_{il}), \\ \pi'_{ij} &= \pi_{ij} + (\pi_{kj} \wedge \pi_{il}), & \pi'_{kl} &= \pi_{kl} + (\pi_{kj} \wedge \pi_{il}), \end{aligned}$$

leaving all other entries untouched. If  $\pi$  satisfies the conditions of (b), then so does  $\pi'$ . By concavity of the function  $\psi$  we get

$$\psi(b_j - a_k) + \psi(b_l - a_i) \geq \psi(b_j - a_i) + \psi(b_l - a_k).$$

Hence the left hand side of (3.2) decreases when we replace  $\pi$  by  $\pi'$ , and we observe that  $a_k < a_i < b_j < b_l$  are not crossed by  $\pi'$ . We say the crossing is repaired.

We systematically repair all crossings as in Section 5 of [14], i.e. in the following order

- picking  $b_j$  from  $\{b_1, \dots, b_n\}$  from left to right,
- given  $b_j$  picking  $a_i < b_j$  from  $\{a_i : 1 \leq i \leq n, a_i < b_j\}$  from right to left,
- given  $a_i < b_j$  picking  $a_k < a_i$  from  $\{a_{i+1}, a_{i+2}, \dots\}$  from right to left,
- given  $a_k < a_i < b_j$  picking  $b_l$  from  $\{b_{j+1}, b_{j+2}, \dots\}$  from left to right.

Just as in Lemma 5.1 of [14] we see that this procedure is well-defined (taking limits in the last two steps) and leads to a matrix  $\pi^*$ , which satisfies

$$\sum_{i=1}^N \sum_{j=1}^{\infty} + \sum_{i=1}^{\infty} \sum_{j=1}^N \pi_{i,j} \psi(b_j - a_i) \geq \sum_{i=1}^N \sum_{j=1}^{\infty} + \sum_{i=1}^{\infty} \sum_{j=1}^N \pi_{i,j}^* \psi(b_j - a_i).$$

Moreover,  $\pi^*$  crosses no four points  $a_k < a_i < b_j < b_l$  with  $a_i, b_j \in \mathcal{E}$  and hence, as in Lemma 5.2 of [14], we get that  $\pi_{ij}^* = \mathbb{1}\{\tau(a_i) = b_j\}$ . Plugging the entries of  $\pi^*$  into the right hand side gives (3.2).  $\square$

We now use Lemma 3.2 to get a continuous inequality. Let  $\mathcal{E} = [b_0, a_0] \subset \mathbb{R}$  be an excursion and  $M = \ell^\mu(\mathcal{E}) > 0$ . Given  $n \in \mathbb{N}$  we pick  $a_1 > a_2 > \dots > a_n$  such that  $\ell^\mu(a_i, a_{i-1}) = \frac{M}{n}$  and  $b_1 < b_2 < \dots < b_n$  such that  $\ell^\nu(b_{j-1}, b_j) = \frac{M}{n}$ , for  $1 \leq i, j \leq n$ , and also such that  $a_n = b_0$  and  $b_n = a_0$ .

Additionally,  $a_i, b_i$ , for  $i > n$ , are chosen in such a way that  $a_i \searrow -\infty$ ,  $b_i \nearrow \infty$ , and

$$\sup_{i=n}^{\infty} (b_{i+1} - b_i), \quad \sup_{j=n}^{\infty} (a_j - a_{j+1}) \longrightarrow 0.$$

Then, if  $\tau_n : \{\dots, a_3, a_2, a_1\} \rightarrow \{b_1, b_2, b_3, \dots\}$  is the (discrete) stable allocation map, in Lemma 3.2 we may choose  $N = n$ . Now suppose a forward-looking and balancing transport rule  $\theta$  is given. We define

$$\pi_{i,j} = \frac{n}{M} \theta((a_i, a_{i-1}], (b_{j-1}, b_j])$$

and note that  $\pi = (\pi_{i,j} : i, j \in \mathbb{N})$  satisfies the conditions of Lemma 3.2 (b). Hence we have

$$\sum_{i=1}^n \sum_{j=1}^{\infty} + \sum_{i=1}^{\infty} \sum_{j=1}^n \pi_{i,j} \psi(b_j - a_i) \geq 2 \sum_{i=1}^n \psi(\tau_n(a_i) - a_i). \quad (3.3)$$

Multiply both sides by  $M/n$  and let  $n$  go to infinity. We will argue that the right and left hand side of (3.3) converge to those of (3.1).

*First* we show convergence of the left hand side of (3.3). Given  $\{a_i\}_{i \in \mathbb{N}}$  and  $\{b_j\}_{j \in \mathbb{N}}$  as constructed above, let  $g_n(a) = a_i$ , if  $a \in (a_i, a_{i-1}]$  and  $h_n(b) = b_j$ , if  $b \in (b_{j-1}, b_j]$ .

**Lemma 3.4.** *For  $\ell^\mu$ -almost every  $a$  we have  $g_n(a) \rightarrow a$ , and for  $\ell^\nu$ -almost every  $b$  we have  $h_n(b) \rightarrow b$ .*

*Proof.* It suffices to prove the first claim. The result is trivial if  $a \leq b_0$ . Given  $\varepsilon > 0$ , for  $\ell^\mu$ -almost every  $a$  with  $b_0 + \varepsilon < a \leq a_0$  there exists  $\eta > 0$  such that  $\ell^\mu(a - \varepsilon, a) \geq \eta$ . Hence, for all  $n > M/\eta$ , there exists  $a_i \in (a - \varepsilon, a)$  which implies  $0 < a - g_n(a) < \varepsilon$ .  $\square$

Note that

$$\frac{M}{n} \sum_{i=1}^n \sum_{j=1}^{\infty} \pi_{i,j} \psi(b_j - a_i) = \iint_{\mathcal{E} \times \mathbb{R}} \psi(h_n(b) - g_n(a)) \theta(da, db).$$

Now we compare the integrand with  $\psi(b - a)$ . Adding and subtracting  $\psi(b - g_n(a))$  and then using the triangle inequality we get

$$\begin{aligned} |\psi(h_n(b) - g_n(a)) - \psi(b - a)| &\leq |\psi(h_n(b) - g_n(a)) - \psi(b - g_n(a))| + |\psi(b - a) - \psi(b - g_n(a))| \\ &\leq \psi(h_n(b) - b) + \psi(a - g_n(a)), \end{aligned}$$

where in the second inequality we used the sub-additivity of  $\psi$  as follows,

$$\begin{aligned} |\psi(h_n(b) - g_n(a)) - \psi(b - g_n(a))| &= \psi(h_n(b) - g_n(a)) - \psi(b - g_n(a)) \\ &= \psi(h_n(b) - b + b - g_n(a)) - \psi(b - g_n(a)) \leq \psi(h_n(b) - b). \end{aligned}$$

Using this estimate on the integrand, we get

$$0 \leq \iint_{\mathcal{E} \times \mathbb{R}} \psi(h_n(b) - g_n(a)) - \psi(b - a) \theta(da, db) \leq \iint_{\mathcal{E} \times \mathbb{R}} \psi(h_n(b) - b) + \psi(a - g_n(a)) \theta(da, db).$$

The integrand on the right is bounded and converges to zero,  $\theta$ -almost everywhere, by Lemma 3.4. Hence the left hand side of (3.3), multiplied by  $\frac{M}{n}$ , converges to the required limit,

$$\iint_{\mathcal{E} \times \mathbb{R}} + \iint_{\mathbb{R} \times \mathcal{E}} \psi(b - a) \theta(da, db).$$

*Second* we show convergence of the right hand side of (3.3). The key to this is the following lemma.

**Lemma 3.5.** *For  $\ell^\mu$ -almost every  $a \in \mathcal{E}$ , we have  $\lim_{n \rightarrow \infty} \tau_n(g_n(a)) = \tau_*(a)$ .*

*Proof.* We define

$$f: [a, \tau_*(a)] \rightarrow [0, \infty), f(x) = \ell^\mu[a, x] - \ell^\nu[a, x],$$

and

$$f_n: [a, \tau_*(a)] \rightarrow \mathbb{R}, f_n(x) = \frac{M}{n} (|\{i: a_i \in [g_n(a), x]\}| - |\{j: b_j \in [g_n(a), x]\}|).$$

The proof is organised into five steps.

**Step 1:**  $|f(x) - f_n(x)| \leq \frac{4M}{n}$ .

*Proof.* Denote  $k_1 = |\{i: a_i \in [g_n(a), x]\}|$  and  $k_2 = |\{j: b_j \in [g_n(a), x]\}|$ , then  $f_n(x) = \frac{M}{n}(k_1 - k_2)$  and

$$\ell^\mu[a, x] = (k_1 - 1) \frac{M}{n} + \ell^\mu[g_n(x), x] - \ell^\mu[g_n(a), a],$$

and hence  $|\ell^\mu[a, x] - k_1 \frac{M}{n}| \leq \frac{2M}{n}$ . Similarly,  $|\ell^\nu[a, x] - k_2 \frac{M}{n}| \leq \frac{2M}{n}$ , which implies the statement.  $\square$

Recall that  $\tau_*(a) = \inf\{x > a: f(x) = 0\}$  and

$$\tau_n(g_n(a)) = \inf\{x > g_n(a): f_n(x) = 0\}.$$

**Step 2:** For  $\ell^\mu$ -almost every  $a \in \mathcal{E}$ , there exists  $\varepsilon_0 > 0$  such that,

$$\begin{aligned} f(a + x) &\geq \frac{5}{6} \ell^\mu[a, a + x] \text{ for all } x \in (0, \varepsilon_0), \text{ and} \\ f(\tau_*(a) - x) &\geq \frac{5}{6} \ell^\nu[\tau_*(a) - x, \tau_*(a)] \text{ for all } x \in (0, \varepsilon_0). \end{aligned}$$

*Proof.* For  $\ell^\mu$ -almost every  $a \in \mathcal{E}$ , there exists  $\varepsilon_0 > 0$  such that,

$$\ell^\nu[a, a + x] \leq \frac{1}{6} \ell^\mu[a, a + x] \text{ for all } x \in (0, \varepsilon_0).$$

see, e.g., Section 1.6, Theorem 3 in [6]. This implies the first property, and the second is analogous.  $\square$

**Step 3:** Let  $G_n := \{a_i : i \in \{1, \dots, n\}, \nexists b_j \in [a_i, a_{i-7}]\}$ . Then, for  $\ell^\mu$ -almost every  $a \in \mathcal{E}$ , there exists arbitrarily large  $n$  with  $g_n(a) \in G_n$ .

*Proof.* Let  $\eta > 0$ . For  $\ell^\mu$ -almost every  $x \in \mathcal{E}$ , there exists  $\delta_0 > 0$  such that  $[x - \delta_0, x + \delta_0] \subset \mathcal{E}$  and

$$\ell^\nu[x - \delta, x + \delta] \leq \eta \ell^\mu[x - \delta, x + \delta], \quad \text{for all } 0 < \delta < \delta_0,$$

see, e.g., Section 1.6, Theorem 3 in [6]. Supposing that, for some  $k \in \mathbb{N}$ ,

$$\frac{(k+2)M}{n} \leq \ell^\mu[x - \delta, x + \delta] \leq \frac{(k+3)M}{n}$$

we have

$$|\{a_i : i \in \{1, \dots, n\}, a_i \in [x - \delta, x + \delta]\}| \geq k + 1$$

and

$$|\{b_j : j \in \{1, \dots, n\}, b_j \in [x - \delta, x + \delta]\}| \leq 1 + \eta(k + 3).$$

By the pigeonhole principle therefore

$$|\{a_i \in [x - \delta, x + \delta] : a_i \notin G_n\}| \leq 7(1 + \eta(k + 3)).$$

Hence, given  $\varepsilon > 0$ , we can find  $\eta$ , and hence  $\delta_0(x) > 0$ , such that

$$\ell^\mu\{a \in [x - \delta, x + \delta] : g_n(a) \notin G_n\} < \frac{\varepsilon}{4M} \ell^\mu[x - \delta, x + \delta], \quad \text{for all } 0 < \delta < \delta_0(x) \text{ and } n \geq n_0(\delta, x).$$

Let

$$\mathcal{E}(\delta) := \{x \in \mathcal{E} : \delta_0(x) > \delta\}.$$

We now fix a global  $\delta > 0$  with  $\ell^\mu(\mathcal{E} \setminus \mathcal{E}(\delta)) < \frac{\varepsilon}{2}$ . We then cover  $\mathcal{E}(\delta)$  by finitely many intervals  $[x_1 - \delta, x_1 + \delta], \dots, [x_m - \delta, x_m + \delta]$  with  $x_i \in \mathcal{E}(\delta)$  such that each of the centres  $x_1, \dots, x_m$  is only contained in one of the intervals. Let  $n_0 := \max_{i=1}^m n_0(\delta, x_i)$ . Then, for all  $n \geq n_0$ , we have

$$\begin{aligned} \ell^\mu\{a \in \mathcal{E} : g_n(a) \notin G_n\} &\leq \sum_{i=1}^m \ell^\mu\{a \in [x_i - \delta, x_i + \delta] : g_n(a) \notin G_n\} + \ell^\mu(\mathcal{E} \setminus \mathcal{E}(\delta)) \\ &\leq \frac{\varepsilon}{4M} \sum_{i=1}^m \ell^\mu[x_i - \delta, x_i + \delta] + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{4M} 2\ell^\mu(\mathcal{E}) + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

The result follows as  $\varepsilon > 0$  was arbitrary.  $\square$

**Step 4:** Using Step 3, for  $\ell^\mu$ -almost every  $a$ , we can choose  $n$  such that  $g_n(a) \in G_n$  and

$$\frac{5M}{n} \leq \min\{f(y) : y \in [a + \varepsilon, \tau_*(a) - \varepsilon]\}.$$

Recall that  $g_n(a) = a_i$ , if  $a \in (a_i, a_{i-1}]$ . In this case we also write  $g_n^{(k)}(a) = a_{i-k}$ .

By Step 1 and choice of  $n$  we have

$$f_n(x) \geq f(x) - \frac{4M}{n} \geq \frac{M}{n} \text{ for all } x \in [a + \varepsilon, \tau_*(a) - \varepsilon].$$

By Step 2, using again Step 1,

$$f_n(x) \geq f(x) - \frac{4M}{n} \geq \frac{5}{6} \ell^\mu[a, g_n^{(7)}(a)] - \frac{4M}{n} \geq \frac{M}{n} \text{ for all } x \in [g_n^{(7)}(a), a + \varepsilon],$$

and the fact that  $g_n(a) \in G_n$  implies that

$$f_n(x) \geq \frac{M}{n} \text{ for all } x \in [g_n(a), g_n^{(7)}(a)].$$

We thus obtain that

$$\tau_n(g_n(a)) = \inf\{x > g_n(a) : f_n(x) = 0\} \geq \tau_*(a) - \varepsilon.$$

**Step 5:** For  $\ell^\mu$ -almost every  $a$ , we have, for sufficiently large  $n$ ,

$$\tau_n(g_n(a)) \leq \tau_*(a) + \varepsilon.$$

*Proof.* Recall that  $h_n(b) = b_j$ , if  $b \in (b_{j-1}, b_j]$ . In this case we also write  $h_n^{(k)}(b) = b_{j+k}$ . We note from Step 1 that  $f_n(\tau_*(a)) \leq \frac{4M}{n}$ . Let  $G'_n := \{b_j : j \in \{1, \dots, n\}, \exists a_i \in [b_j, b_{j+5}]\}$ . Then, as in Step 3, for  $\ell^\mu$ -almost every  $a$ , there exists arbitrarily large  $n$  with  $h_n(\tau_*(a)) \in G'_n$ . We infer that  $\tau_n(g_n(a)) \leq h_n^{(5)}(\tau_*(a)) \leq \tau_*(a) + \varepsilon$ , if  $n$  is large enough.  $\square$

The result follows by combining Steps 4 and 5, as  $\varepsilon > 0$  was arbitrarily small.  $\square$

To conclude we note that, using Lemmas 3.4 and 3.5,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{M}{n} \sum_{i=1}^n \psi(\tau_n(a_i) - a_i) &= \lim_{n \rightarrow \infty} \int_{\mathcal{E}} \psi(\tau_n(g_n(a)) - g_n(a)) \ell^\mu(da) \\ &= \int_{\mathcal{E}} \lim_{n \rightarrow \infty} \psi(\tau_n(g_n(a)) - g_n(a)) \ell^\mu(da) = \int \psi(\tau_*(a) - a) \ell^\mu(da), \end{aligned}$$

by bounded convergence.

**4. Ergodicity.** Denote by  $\mathbb{P}^{(\mu)} = \int \mathbb{P}_x \mu(dx)$  the law of two-sided Brownian motion with  $\omega_0 \sim \mu$  or, in other words, the push-forward of  $\mathbb{P}_\mu$  under  $(B_t : t \in \mathbb{R})$ . This is a measure on the path space  $(\Omega, \mathcal{A})$  introduced at the beginning of Section 2. For  $r \in \mathbb{R}$  define  $S^r$  to be the generalized inverse of the local time  $\ell^\mu$ , that is

$$S^r := \begin{cases} \sup\{t \geq 0 : \ell^\mu[0, t] = r\}, & \text{if } r \geq 0, \\ \sup\{t < 0 : \ell^\mu[t, 0] = -r\}, & \text{if } r < 0. \end{cases} \quad (4.1)$$

Then  $\mathbb{P}^{(\mu)}$  is invariant under the shifts  $S^r : (\omega_s) \mapsto (\omega_{S^{r+s}})$ , for every  $r \in \mathbb{R}$ , by Theorem 3.4 in [10]. In order to show that the family of shifts  $(S^r)$  is ergodic, we need to show that any invariant set  $A$  is trivial, i.e.  $\mathbb{P}^{(\mu)}(A) \in \{0, 1\}$ . We follow a classical approximation approach. By [5, Appendix A.3] there exist sets  $A_r \in \sigma(\omega_s : s \in [S^{-r}, S^r])$ ,  $r > 0$ , such that  $\mathbb{P}^{(\mu)}(A \Delta A_r) \rightarrow 0$  as  $r \rightarrow \infty$ , where  $\Delta$  denotes the symmetric difference of the two sets. As  $A$  is invariant under the shift  $S^{2r}$  we get

$$\mathbb{P}^{(\mu)}(A \Delta S^{2r} A_r) = \mathbb{P}^{(\mu)}(S^{2r} A \Delta S^{2r} A_r) = \mathbb{P}^{(\mu)}(S^{2r}(A \Delta A_r)) = \mathbb{P}^{(\mu)}(A \Delta A_r) \rightarrow 0.$$

Hence there exists  $r_n \nearrow \infty$  such that

$$\sum_{n=1}^{\infty} \mathbb{P}^{(\mu)}(A \Delta S^{2r_n} A_{r_n}) < \infty,$$

and  $A \setminus \bigcap_{s>0} \bigcup_{r_n>s} S^{2r_n} A_{r_n}$  and  $\bigcap_{s>0} \bigcup_{r_n>s} S^{2r_n} A_{r_n} \setminus A$  are  $\mathbb{P}^{(\mu)}$ -nullsets. This implies

$$\mathbb{P}^{(\mu)}(A) = \mathbb{P}^{(\mu)}\left(\bigcap_{s>0} \bigcup_{r_n>s} S^{2r_n} A_{r_n}\right) \in \{0, 1\},$$

using that the latter event is a tail event and hence trivial, see e.g. [13, Theorem 2.9].

**Lemma 4.1.** *Let  $T \geq 0$  be an unbiased shift and  $\theta$  be the associated transport rule. Furthermore let  $\psi : [0, \infty) \rightarrow [0, \infty)$  be concave with  $\psi(0) = 0$ . Let  $S_{\mu+\nu}^r$  be the generalized inverse of the local time  $\ell^{\mu+\nu}$ . Then,  $\mathbb{P}^{(\mu+\nu)}$ -almost surely,*

$$\lim_{r \rightarrow \infty} \frac{1}{r} \iint_0^{S_{\mu+\nu}^r} \psi(t-s) \theta(ds, dt) = \lim_{r \rightarrow \infty} \frac{1}{r} \iint_{S_{\mu+\nu}^{-r}}^0 \psi(t-s) \theta(ds, dt) = \frac{1}{2} \mathbb{E}_\mu \psi(T),$$

and,

$$\lim_{r \rightarrow \infty} \frac{1}{r} \int_0^{S_{\mu+\nu}^r} \int \psi(t-s) \theta(ds, dt) = \lim_{r \rightarrow \infty} \frac{1}{r} \int_{S_{\mu+\nu}^{-r}}^0 \int \psi(t-s) \theta(ds, dt) = \frac{1}{2} \mathbb{E}_\mu \psi(T).$$

*Proof.* By the ergodic theorem,  $\mathbb{P}^{(\mu+\nu)}$ -almost surely,

$$\begin{aligned} \lim_{r \rightarrow \infty} \frac{1}{r} \iint_0^{S_{\mu+\nu}^r} \psi(t-s) \theta(ds, dt) &= \lim_{r \rightarrow \infty} \frac{1}{r} \int_0^{S_{\mu+\nu}^r} \left( \int \psi(t-s) \theta(s, dt) \right) \ell^{\mu+\nu}(ds) \\ &= \lim_{r \rightarrow \infty} \frac{1}{r} \int_0^r \left( \int \psi(t - S_{\mu+\nu}^s) \theta(S_{\mu+\nu}^s, dt) \right) ds \\ &= \frac{1}{2} \mathbb{E}^{(\mu+\nu)} \int \psi(t) \theta(0, dt) = \frac{1}{2} \mathbb{E}_\mu \psi(T). \end{aligned}$$

Similarly, we obtain  $\mathbb{P}^{(\mu+\nu)}$ -almost surely,

$$\begin{aligned} \lim_{r \rightarrow \infty} \frac{1}{r} \int_0^{S_{\mu+\nu}^r} \int \psi(t-s) \theta(ds, dt) &= \frac{1}{2} \mathbb{E}^{(\mu+\nu)} \int \psi(-s) \theta(ds, 0) \\ &= \frac{1}{2} \mathbb{E}^{(\nu)} \int \psi(-s) \theta(ds, 0). \end{aligned}$$

Using the generalized Campbell formula, see [10, (2.4)], we get

$$\begin{aligned} \mathbb{E}^{(\nu)} \int \psi(-s) \theta(ds, 0) &= \mathbb{E}^{(\nu)} \int \left( \int \psi(-s) \mathbf{1}\{s+r \in [0, 1]\} \theta_\omega(ds, 0) \right) dr \\ &= \mathbb{E} \int \left( \int \psi(-s) \mathbf{1}\{s+r \in [0, 1]\} \theta_{r\omega}(ds, 0) \right) \ell^\nu(dr) \\ &= \mathbb{E} \iint \mathbf{1}\{t \in [0, 1]\} \psi(r-t) \theta_\omega(dt, r) \ell^\nu(dr), \end{aligned}$$

where in the last equation we used the shift-invariance of  $\theta$ . Using the balancing property first and then the generalized Campbell formula again this equals

$$\mathbb{E} \iint \mathbf{1}\{t \in [0, 1]\} \psi(r-t) \theta_\omega(t, dr) \ell^\mu(dt) = \mathbb{E}^{(\mu)} \int \psi(t) \theta(0, dt) = \mathbb{E}_\mu \psi(T).$$

The claims about backward time follow in the same manner.  $\square$

*Proof of Theorem 1.* Define for  $u > 0$  the stopping times

$$\begin{aligned} \rho(u) &= \inf\{t \geq 0 : \ell^\mu([0, t]) - \ell^\nu([0, t]) = -u\}, \\ \sigma(u) &= \sup\{t \leq 0 : -\ell^\mu([t, 0]) + \ell^\nu([t, 0]) = -u\}. \end{aligned}$$

We have  $\rho(u) \nearrow \infty$  and  $\sigma(u) \searrow -\infty$ , as  $u \rightarrow \infty$ , and, for all  $u > 0$ , the interval  $[\sigma(u), \rho(u)]$  forms an excursion. Hence, by Proposition 3.1,

$$\int_{\sigma(u)}^{\rho(u)} \int \psi(t-s) \theta(ds, dt) + \int \int_{\sigma(u)}^{\rho(u)} \psi(t-s) \theta(ds, dt) \geq 2 \int_{\sigma(u)}^{\rho(u)} \psi(\tau_*(s) - s) \ell^\mu(ds).$$

Applying Lemma 4.1 the left hand side is asymptotically equal to  $\frac{1}{2} \ell^{\mu+\nu}([\sigma(u), \rho(u)]) \mathbb{E}_\mu \psi(T)$ , and the right hand side to  $\frac{1}{2} \ell^{\mu+\nu}([\sigma(u), \rho(u)]) \mathbb{E}_\mu \psi(T_*)$ , which concludes the proof.  $\square$

**Acknowledgements:** This paper is part of the second author's PhD project. We would like to thank the referee for their careful reading of the manuscript.

**References.**

- [1] M. Beiglböck, A. M. G. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Inventiones Math.*, to appear, 2016.
- [2] J. Bertoin and Y. Le Jan. Representation of measures by balayage from a regular recurrent point. *Ann. Probab.*, 20:538–548, 1992.
- [3] A. M. G. Cox and D. G. Hobson. An optimal Skorokhod embedding for diffusions. *Stochastic Process. Appl.*, 111:17–39, 2004.
- [4] A. M. G. Cox and D. G. Hobson. A unifying class of Skorokhod embeddings: connecting the Azéma-Yor and Vallois embeddings. *Bernoulli*, 13:114–130, 2007.
- [5] R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.
- [6] L. C. Evans and R. F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [7] A. E. Holroyd and T. M. Liggett. How to find an extra head: optimal random shifts of Bernoulli and Poisson random fields. *Ann. Probab.*, 29:1405–1425, 2001.
- [8] A. E. Holroyd, R. Pemantle, Y. Peres, and O. Schramm. Poisson matching. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45:266–287, 2009.
- [9] A. E. Holroyd and Y. Peres. Extra heads and invariant allocations. *Ann. Probab.*, 33:31–52, 2005.
- [10] G. Last, P. Mörters, and H. Thorisson. Unbiased shifts of Brownian motion. *Ann. Probab.*, 42:431–463, 2014.
- [11] G. Last and H. Thorisson. Invariant transports of stationary random measures and mass-stationarity. *Ann. Probab.*, 37:790–813, 2009.
- [12] T. M. Liggett. Tagged particle distributions or how to choose a head at random. In *In and out of equilibrium (Mambucaba, 2000)*, volume 51 of *Progr. Probab.*, pages 133–162. Birkhäuser, Boston, MA, 2002.
- [13] P. Mörters and Y. Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010.
- [14] P. Mörters and I. Redl. Skorokhod embeddings for two-sided Markov chains. *Probability Theory and Related Fields*, 165:483–508, 2016.
- [15] H. Thorisson. *Coupling, Stationarity, and Regeneration*. Probability and its Applications (New York). Springer-Verlag, New York, 2000.

Peter Mörters and István Redl  
Department of Mathematical Sciences  
University of Bath  
Bath BA2 7AY  
United Kingdom