FASTEST COUPLING OF RANDOM WALKS

L C G Rogers

University of Bath

This draft: April 1997

Tel: +44 1225 826224
Fax: +44 1225 826492
Email: lcgr@maths.bath.ac.uk

Abstract

We describe a new coupling of one-dimensional random walks which tries to control the coupling by keeping the separation of the two random walks of constant sign. It turns out that among such monotone couplings there is an optimal one-step coupling which maximises the second moment of the difference (assuming this is finite), and this coupling is 'fast' in the sense that for a random walk with a unimodal step distribution the coupling time achieved by using the new coupling at each step is stochastically no larger than any other coupling. We apply this to the case of symmetric unimodal distributions.

Abbreviated title: FASTEST COUPLING OF RANDOM WALKS

Key words and phrases: Coupling, random walk, unimodal, Cauchy, monotone coupling, maximal coupling.

AMS 1991 Subject Classifications: Primary 60J15

Research supported in part by EPSRC grant number GR/J97281

FASTEST COUPLING OF RANDOM WALKS

L C G Rogers

University of Bath

1. Introduction. Over the last two decades or so, the coupling method has established itself as a powerful technique for proving convergence results, and has been applied to a large range of processes, such as random walks, renewal processes, particle systems; the beautiful survey by Lindvall [5] presents the main ideas and many typical applications. The object of this paper is to study 'monotone coupling' of real-valued random walks; by this we mean that if $(S_n)_{n\geq 0}$ and $(S'_n)_{n\geq 0}$ are two real-valued random walks with common step distribution F but different initial conditions, $S_0 < S'_0$, we shall be considering couplings (that is, joint laws for the two processes S and S') with the property that

$$S_n \leq S'_n \text{ for all } n.$$

A coupling is called *successful* if the coupling time $T \equiv \inf\{n : S_n = S'_n\}$ is almost surely finite, and we may and shall assume that $S_n = S'_n$ for all $n \geq T$.

The advantage of looking at monotone couplings is that the difference process $S'_n - S_n$ is a non-negative martingale[1], and this guarantees the almost-sure convergence of the difference as $n \to \infty$. This does not (of course) ensure successful coupling, but helps in that direction.

A coupling may be successful, but is it good? If $\mathcal{L}(\theta_n S)$ denotes the law of the process $(S_n, S_{n+1}, S_{n+2}, \ldots)$, the coupling inequality (see I.2 in [5]) gives the bound on the total-variation norm

$$||\mathcal{L}(\theta_n S) - \mathcal{L}(\theta_n S')|| \leq 2\mathbb{P}(T > n),$$

and for a coupling to be good, we would want the bound to be tight. In fact, it is known (see Goldstein [2], Thorisson [8]) that one *can* construct a coupling for which the coupling inequality holds with equality for all n, but the Goldstein-Thorisson construction is quite abstract. If we call such a coupling a *fastest* coupling[2], then the main result of this paper is an explicit construction of a fastest coupling in the special case where the step distribution is unimodal.

In more detail, in Section 2 we shall consider just a single step of the random walks S and S', assuming that $S_0 = 0$ and $S'_0 = a > 0$. We shall then construct a pair (X, Y) such that

(1.i) $X \leq Y$ almost surely;

(1.ii) X has law F and Y has law $F_a \equiv \delta_a * F$;

(1.iii) $\mathbb{E}\ \varphi(Y - X)$ is maximal for any non-negative decreasing convex function φ.

To amplify, we aim to maximise $\mathbb{E}\ \varphi(Y - X)$ over choice of joint laws of (X, Y) which satisfy (1.i) and (1.ii). It is not obvious that there should be a unique joint law which achieves

[1] assuming that the step distribution has finite first moment

[2] also known as a *maximal* coupling

this maximum whatever convex decreasing $\varphi \geq 0$ we take, but this is what happens, as we shall prove in Section 2. The joint law of (X, Y) is given explicitly in term of the marginals; see (4). Our problem (1) is a problem of Monge-Kantorovich type; for surveys, see for example, [2] or [7]. The result proved here appears to be new.

The motivation for trying to achieve (1.iii) is that we would like to maximise $\mathbb{E}[(X - Y)^2]$, so as to create a difference process with the largest possible quadratic variation, and so hopefully with a short coupling time. If F has finite second moment, then the joint law which optimises (1.i)–(1.iii) also maximises the second moment of the difference $X - Y$, but the problem is well posed without assuming finite second moments, and the generality of (1.iii) is needed for the next stage of the paper.

The optimality property of Section 2 is myopic, in that for a given step distribution we seek to achieve a 'best' coupling of the single step about to be made, but it turns out that *in the case where F is unimodal*, this myopic strategy is globally optimal; the coupling is fastest. This is proved in Section 3.

The construction of Sections 2 and 3 is concrete, but not particularly easy to visualise. However, as we show in Section 4, in the case of *symmetric* unimodal step distributions, the construction is much more explicit, and we are able to derive a simple expression for the total-variation distance:

$$\begin{aligned} ||\mathcal{L}(\theta_n S) - \mathcal{L}(\theta_n S')|| &= 2\mathbb{P}(T > n) \\ &= 2\mathbb{P}(|S_n| < a/2). \end{aligned}$$

We are then able to exhibit examples of random walks which couple very rapidly indeed; while the best rate of convergence one can achieve in general is $O(n^{-1/2})$ (see, for example, Lindvall & Rogers [6]), by taking symmetric unimodal distributions with much fatter tails we can get very fast couplings, even couplings for which the coupling time has finite moments!

Acknowledgement. It is a pleasure to thank Torgny Lindvall for stimulating discussions on the issues discussed in this paper, and other topics. Anyone reading this article will join me in thanking also an exceptionally diligent and perceptive referee, whose careful reading of the manuscript resulted in many improvements.

2. Monotone coupling of a single step. Here is the main result of this section.

THEOREM 1. *Suppose that F and G are two distribution functions, such that G is stochastically larger than F:*

$$G(x) \leq F(x) \tag{2}$$

for all x, and suppose that $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ is strictly convex and decreases to 0. The problem of finding a probability measure P on $\mathbb{R}^2$ such that:

(3.i) P is concentrated on $\{(x, y) : x \leq y\}$;

(3.ii) the marginals of P are F and G;

(3.iii) the payoff

$$\int\int \varphi(y-x)P(dx,dy)$$

is maximised

is solved by taking for $x \le y$

$$
\begin{aligned}
P(x,y) &\equiv P((-\infty,x]\times(-\infty,y]) \\
&= \Gamma(x,y) \\
&\equiv \sup_{s\le x}\ \sup_{s\le v\le y}\ (G(v)-F(v)+F(s))\,, \qquad (4)
\end{aligned}
$$

and $P(x,y) = P(y,y)$ for $x > y$.

Remarks. (i) The condition (2), that G is stochastically larger than F, will certainly be satisfied if $G = F_a \equiv \delta_a * F$.

(ii) It is tempting to conjecture that the Theorem might still be true if we did not insist on (3.i), if we extend φ to $(-\infty,0)$ symmetrically. If this were the case, optimality would force $X \le Y$ P-almost surely. However, without (3.i) the result collapses, and in Figure 1 we give a simple example concentrated on $\{0,1,\dots,12\}$ for which the conjecture fails. The positive integers shown in the grid need to be normalised to become a probability distribution, but modulo this trivial transformation, what is shown in Figure 1 is the optimal monotone coupling, as at (4). However, if $\varphi(1)+\varphi(7) > 2\varphi(3)$, as is perfectly possible, then we can improve the payoff by increasing the probability on the starred squares (6,5) and (2,7), and reducing the probability on (2,5) and (6,9) to zero. This shows also that without the monotonicity restriction the optimal joint law may be different for different φ.

(iii) In the example of Figure 1, notice also that the law G is in fact F_1, the law of the first marginal shifted up by 1.

Proof. To begin with, the payoff can be re-expressed as

$$
\begin{aligned}
\int\int \varphi(y-x)P(dx,dy) &= \int\int \left(\int_{t=y-x}^{\infty} -\varphi'(t)dt\right) P(dx,dy) \\
&= \int\int \left(\int_{t=y-x}^{\infty}\int_{s=t}^{\infty} \varphi''(ds)dt\right) P(dx,dy) \\
&= \int_0^{\infty} \varphi''(ds) \int_{-\infty}^{\infty} \mathbb{P}[Y\le u, X > u-s]\,du. \qquad (5)
\end{aligned}
$$

Thus the problem can be seen as one of finding good upper bounds for $\mathbb{P}[Y \le u, X > u-s]$ for pairs of random variables (X,Y) with marginals F and G, and for which $X \le Y$. Take $x < y$ and observe that if P satisfies (3.i) and (3.ii) then

$$
\begin{aligned}
\mathbb{P}(Y\le y, X > x) &\le F(y) - F(x) \\
\mathbb{P}(Y\le y, X > x) &\le G(y) - G(x),
\end{aligned}
$$

so that

$$P(x,y) = G(y) - \mathbb{P}(Y \le y, X > x) \tag{6}$$
$$\ge G(y) - (F(y) - F(x)) \wedge (G(y) - G(x))$$
$$= G(x) \vee (F(x) + G(y) - F(y)).$$

Since $P(\cdot\,,\cdot)$ is increasing in both arguments, we must have that

$$P(x,y) \ge h(x,y) \equiv \sup_{x \le v \le y} \{G(v) - F(v) + F(x)\} \tag{7}$$

The function h is clearly increasing in its second argument, but not always increasing in its first; we conclude that any P satisfying (3.i) and (3.ii) must satisfy the inequality

$$P(x,y) \ge \Gamma(x,y) \equiv \sup_{s \le x} h(s,y), \tag{8}$$

for all $x \le y$, rewriting the definition of Γ more briefly. Combining (8), (6) and (5), we deduce that for any convex function φ decreasing to 0 and P satisfying (3.i) and (3.ii)

$$\mathbb{E}\, \varphi(Y - X) \le \int_0^\infty \varphi''(ds) \int_{-\infty}^\infty (G(u) - \Gamma(u - s, u))du.$$

The proof will be completed by showing that Γ *is* a distribution function satisfying (3.i) and (3.ii).

Firstly we show that the marginals of Γ are correct. Noting that for $x \le y$ we have $h(x,y) \le G(y)$, we have that

$$\lim_{x \to \infty} \Gamma(x,y) = \Gamma(y,y)$$
$$= \sup_{s \le y} h(s,y)$$
$$\le G(y)$$

and on the other hand $h(s,y) \ge G(s)$ for all $s \le y$, so

$$\Gamma(y,y) = \sup_{s \le y} h(s,y) \ge G(y),$$

and the second marginal is seen to be G, as it should be. As for the first marginal, we see that (because $G \le F$)

$$h(x,y) \equiv \sup_{x \le v \le y} \{G(v) - F(v) + F(x)\} \le F(x)$$

for $x \le y$, and letting $y \uparrow \infty$ we see that

$$\lim_{y \uparrow \infty} h(x,y) = F(x),$$

from which immediately $\lim_{y\uparrow\infty}\Gamma(x,y) = F(x)$, confirming the first marginal to be what it should be.

All that remains, therefore, is to prove that the function Γ actually *is* a distribution function. To spell it out, this requires us to show that for $x < x' \le y < y'$,

$$\Gamma(x',y') - \Gamma(x,y') - \Gamma(x',y) + \Gamma(x,y) \ge 0. \tag{9}$$

In fact, we really should show this inequality for *any* pairs $x < x'$ and $y < y'$, but since Γ does not charge $\{(x,y) : x > y\}$, it is enough to prove (9) for the more restricted range $x < x' \le y < y'$.

If we define for $x \le y$

$$m(x,y) \equiv \inf_{v\in[x,y]} \left(F(v) - G(v)\right), \tag{10}$$

then we see that m is non-negative, increasing in its first argument, and decreasing in its second argument; and for $s \le y \le y'$,

$$m(s,y') = m(s,y) \wedge m(y,y').$$

Moreover, we can express h in terms of m as $h(x,y) = F(x) - m(x,y)$, so that $\Gamma(x,y) = \sup_{s\le x}(F(s) - m(s,y))$, and for $u \le y \le y'$ we have

$$\begin{aligned}
\Gamma(u,y') &= \sup_{s\le u}\{F(s) - (m(s,y) \wedge m(y,y'))\} \\
&= \sup_{s\le u}\{F(s) - m(s,y) + (m(s,y) - m(y,y'))^+\}, \\
\Gamma(u,y) &= \sup_{s\le u}\{F(s) - m(s,y)\}.
\end{aligned}$$

Thus if we hold fixed y and y' with $y < y'$ and let $s^* \equiv \inf\{s \le y : m(s,y) > m(y,y')\}$, we shall have that $\Gamma(s,y') - \Gamma(s,y)$ will be zero for $s < s^*$, and will be increasing for $s^* \le s \le y$. To see this, observe that the functions $s \mapsto F(s) - m(s,y)$ and $s \mapsto F(s) - m(s,y')$ differ by a non-negative incresaing function. Any point of increase $s > s^*$ of $\Gamma(\cdot,y)$ is automatically a point of increase of $\Gamma(\cdot,y')$, and the increase of $\Gamma(\cdot,y')$ over a subinterval of $[s^*,y]$ is at least as large as the increase of $\Gamma(\cdot,y)$. The inequality (9) follows.

3. Fastest coupling for unimodal distributions. In this section, we shall suppose that the step distribution F is *unimodal*, which is to say that there is some real value (without loss of generality, 0) such that to the right of that value F is concave, and to the left of that value F is convex. We are going to use the monotone coupling of two random walks with step distribution F inspired by Theorem 1, which we shall call the convex-monotone (CM) coupling. In more detail, if $S_0 = 0$ and $S'_0 = a > 0$, we shall make the joint distribution of (S_1, S'_1) according to the recipe (4), when we use $G = F_a$. If the realised value of (S_1, S'_1) is (b,c) we then make the joint law of (S_2, S'_2) according to the recipe (4), taking the first marginal to be F_b and the second marginal to be F_c. Subsequent steps of the process are generated analogously.

The main result of this section will be the following.

THEOREM 2. *Assume that the step distribution F is unimodal. Consider two couplings $(S_n, S'_n)_{n\geq 0}$ and $(\tilde{S}_n, \tilde{S}'_n)_{n\geq 0}$, where $S_0 = \tilde{S}_0 = 0$ and $S'_0 = \tilde{S}'_0 = a > 0$. Suppose that the coupling (S, S') is the CM-coupling constructed in Theorem 1, and that $(\tilde{S}, \tilde{S}')$ is any coupling. If*

$$T \equiv \inf\{n \geq 0 : S_n = S'_n\}, \quad \tilde{T} \equiv \inf\{n \geq 0 : \tilde{S}_n = \tilde{S}'_n\},$$

then for any $n \geq 0$,

$$\mathbb{P}(T \geq n) \leq \mathbb{P}(\tilde{T} \geq n). \tag{11}$$

Remarks. Theorem 2 tells us that if the step distribution is unimodal, then the coupling time for the CM-coupling is stochastically smallest among all couplings of the random walks, even without restricting the coupling to be monotone.

Proof. In fact, what we shall prove is stronger than the statement of Theorem 2.

To begin with, let us fix $\lambda > 0$ and define for each $n \in \mathbb{Z}^+$ the function

$$V_n(a) \equiv \sup\{\mathbb{E}\exp(-\lambda|S'_n - S_n|) : S'_0 = a, S_0 = 0\},$$

the sup being taken over all joint laws of (S, S') with $S'_0 = a$, $S_0 = 0$. The functions V_n satisfy the dynamic programming equation

$$V_{n+1}(a) = \sup\ \mathbb{E}\ [V_n(S'_1 - S_1)|S_0 = 0, S'_0 = a]\,. \tag{12}$$

It is clear that the function V_n is symmetric about 0, and is bounded between 0 and 1. The key to the proof is the following result, which is of independent interest.

LEMMA 1. *Fix $a > 0$. Suppose that $\varphi : \mathbb{R} \to \mathbb{R}^+$ is symmetric, convex and decreasing in $\mathbb{R}^+$, and that F is unimodal. Then the problem*

$$\max\{\ \mathbb{E}\varphi(Y - X)\ : X \sim F,\ Y \sim F_a\}$$

is solved by the CM-coupling of F and F_a.

Remarks. (i) The point of Lemma 1 is that even without requiring that $X \leq Y$, the optimal joint law will have this property, provided F is unimodal. The earlier example of Figure 1 shows that this is not true in general without unimodality.

(ii) The proof provides a pretty description of the optimal joint law of (X, Y). As a particular special case, if F is symmetric about 0 and unimodal, with density f, we have

$$\mathbb{P}(X \in dx, Y = X) = [f(x) \wedge f(x - a)]dx \quad \text{for } x \in \mathbb{R};$$
$$\mathbb{P}(X \in dx, Y = a - X) = [f(x) - f(x - a)]dx \quad \text{for } x \leq a/2.$$

Proof of Lemma 1. As in the proof of Theorem 1, with P denoting the joint law of (X,Y), we estimate

$$\begin{aligned}
\mathbb{E}\,\varphi(Y-X) &= \int\int_{\{x>y\}} P(dx,dy)\int_{-\infty}^{y-x} dt \int_{-\infty}^{t}\varphi''(ds) \\
&\qquad + \int\int_{\{x\le y\}} P(dx,dy)\int_{y-x}^{\infty} dt\int_{t}^{\infty}\varphi''(ds) \\
&= \int_0^\infty \varphi''(ds)\{\int_0^s dt\,[\mathbb{P}(X-t<Y<X)+\mathbb{P}(X\le Y\le X+t)]\} \\
&= \int_0^\infty \varphi''(ds)\{\mathbb{E}\left[(s-X+Y)^+;Y<X\right]+\mathbb{E}\left[(s-Y+X)^+;Y\ge X\right]\} \\
&= \int_0^\infty \varphi''(ds)\int_{-\infty}^{\infty} dt\,(\mathbb{P}[t-s<Y<X\le t]+\mathbb{P}[t-s<X\le Y\le t]) \\
&= \int_0^\infty \varphi''(ds)\int_{-\infty}^{\infty} dt\;\mathbb{P}[(X,Y)\in(t-s,t]^2] \qquad (13).
\end{aligned}$$

Thus we want to make $\mathbb{P}[(X,Y)\in(t-s,t]^2]$ as large as possible for each s and t, subject to the constraints of the marginals of (X,Y) being F and F_a. Now it is clear that

$$\mathbb{P}[(X,Y)\in(t-s,t]^2]\le (F(t)-F(t-s))\wedge(F_a(t)-F_a(t-s)), \tag{14}$$

and we shall now show that if F is unimodal then the CM-coupling achieves this obvious bound, and so must be optimal.

To prove this, we need to observe that if F is unimodal, then the function $x\mapsto F(x)-F(x-a)$ is non-negative, and is increasing to the left of some point, decreasing to the right of that point. (Think of the slope of the chord from $(x-a,F(x-a))$ to $(x,F(x))$ as x increases!) Thus for $x\le y$

$$\begin{aligned}
\Gamma_a(x,y) &\equiv \sup_{s\le x}\sup_{s\le v\le y}(F_a(v)-F(v)+F(s)) \\
&= \sup_{s\le x}\left(F(s)-\inf_{s\le v\le y}\{F(v)-F_a(v)\}\right) \\
&= \sup_{s\le x}(F(s)-\{F(s)-F_a(s)\}\wedge\{F(y)-F_a(y)\}) \\
&= \sup_{s\le x}(F_a(s)\vee\{F_a(y)-F(y)+F(s)\}) \\
&= F_a(x)\vee\{F_a(y)-F(y)+F(x)\}.
\end{aligned}$$

Now apply this: we obtain for the measure μ with distribution function Γ_a

$$\begin{aligned}
\mu((x,y]\times(x,y]) &= \Gamma_a(y,y)-\Gamma_a(x,y)-\Gamma_a(y,x)+\Gamma_a(x,x) \\
&= F_a(y)-\Gamma_a(x,y) \\
&= \{F_a(y)-F_a(x)\}\wedge\{F(y)-F(x)\}, \qquad (15)
\end{aligned}$$

as required. Thus the CM-coupling is the optimal coupling without monotonicity constraint, assuming that F is unimodal.

To conclude the proof of Theorem 2, define for $c > 0$ the function $\varphi_c(x) = (c - |x|)^+$, which is symmetric, and convex decreasing in $\mathbb{R}^+$.

LEMMA 2. *If we define for* $a, c > 0$

$$\psi(a,c) \equiv \sup\{\mathbb{E}\,\varphi_c(Y - X) \ : X \sim F, Y \sim F_a\}, \tag{16}$$

then for any $a, c > 0$ *we have*

$$\psi(a,c) + a = \psi(c,a) + c. \tag{17}$$

Before we prove Lemma 2, let us see how Theorem 2 follows from it. From (16) it is clear that ψ is convex in its second argument, since the optimal joint law is the same whatever c. From Lemma 2, therefore, ψ is also convex in its *first* argument. Now

$$\begin{aligned} V_1(a) &\equiv \sup\{\mathbb{E}\,\exp(-\lambda|Y - X|)\,; Y \sim F_a, X \sim F\} \\ &= \int\int \Gamma_a(dx,dy)\ \exp(-\lambda|y-x|), \quad \text{by Lemma 1;} \\ &= \int\int \Gamma_a(dx,dy)\ \int_0^\infty \lambda^2 e^{-\lambda c}(c - |y-x|)^+\, dc \\ &= \int_0^\infty \lambda^2 e^{-\lambda c}\psi(a,c)\, dc \end{aligned}$$

is thus a convex function of $a > 0$, and is plainly symmetric in a. Similarly, if we suppose that $V_n(\cdot)$ is symmetric, and convex in $\mathbb{R}^+$, then it is clear that V_n must be decreasing in $\mathbb{R}^+$, since it is bounded and convex. Next, for $a > 0$ we have

$$\begin{aligned} V_{n+1}(a) &\equiv \sup\{\mathbb{E}\,V_n(Y - X)\,; X \sim F, Y \sim F_a\} \\ &= \int\int \Gamma_a(dx,dy)\ V_n(|y-x|) \\ &= \int\int \Gamma_a(dx,dy)\ \int_0^\infty V_n''(dc)\,(c - |y-x|)^+ \\ &= \int_0^\infty V_n''(dc)\,\psi(a,c) \end{aligned}$$

is convex in $a > 0$, and by induction V_n is convex in $\mathbb{R}^+$ whatever n, and the optimal joint law of the steps of the random walk is given by the myopic policy.

Proof of Lemma 2. We prove the result firstly under the assumption that the distribution F is strictly unimodal, that is, F is strictly convex in $(-\infty, 0)$ and strictly concave in $(0, \infty)$. The choice of 0 as the mode is of course an unimportant convention.

For each $a > 0$, there is for each $t > 0$ a unique solution $x = \xi(t,a)$ to the equation

$$F(x) - F(x-a) = F(x+t) - F(x+t-a),$$

because the function $x \mapsto F(x) - F(x-a)$ is strictly increasing up to some point, then strictly decreasing. It is easy to verify that

$$\xi(a,t) + a = \xi(t,a) + t \tag{18}$$

for all $a, t > 0$.

Now from (13) we have

$$\mathbb{E}\, \varphi_c(Y - X) = \int_{-\infty}^{\infty} dt\; \mathbb{P}[(X,Y) \in (t-c,t]^2],$$

and from (15) we see that when the optimal joint law Γ_a is used this is equal to

$$\begin{aligned}
\psi(a,c) &\equiv \mathbb{E}\, \varphi_c(Y-X) \\
&= \int_{-\infty}^{\infty} (F(s) - F(s-c)) \wedge (F_a(s) - F_a(s-c))\; ds \\
&= \int_{-\infty}^{\xi(a,c)+a} (F(s-a) - F(s-a-c))\; ds + \int_{\xi(a,c)+a}^{\infty} (\bar{F}(s-c) - \bar{F}(s))\; ds, \\
&\qquad\qquad \text{where } \bar{F} \equiv 1 - F; \\
&= \int_{\xi(a,c)-c}^{\xi(a,c)} F(s)\; ds + \int_{\xi(a,c)-c+a}^{\xi(a,c)+a} \bar{F}(s)\; ds \\
&= \int_{\xi(c,a)-a}^{\xi(c,a)+c-a} F(s)\; ds + \int_{\xi(c,a)}^{\xi(c,a)+c} \bar{F}(s)\; ds \\
&= \int_{\xi(c,a)-a}^{\xi(c,a)} F(s)\; ds + \int_{\xi(c,a)-a+c}^{\xi(c,a)+c} \bar{F}(s)\; ds \\
&\qquad\qquad + \int_{\xi(c,a)}^{\xi(c,a)+c-a} F(s)\; ds + \int_{\xi(c,a)}^{\xi(c,a)+c-a} \bar{F}(s)\; ds \\
&= \psi(c,a) + c - a.
\end{aligned}$$

4. Symmetric unimodal steps. Throughout this section, we assume that the step distribution F is symmetric and unimodal, with a density f. The case where F has an atom at 0 is an easy extension which we leave to the interested reader, in order to keep notation simple. We need a simple result firstly about convolutions of symmetric unimodal distributions.

LEMMA 3. *Suppose that F and G are symmetric unimodal distributions with densities f and g respectively. Then $F * G$ is again symmetric unimodal.*

Proof. In the case where F is uniform on $[-b, b]$ we have that the density of $F * G$ is

$$x \mapsto \int_{x-b}^{x+b} g(y) \, \frac{dy}{2b},$$

which is clearly unimodal and symmetric. But the general symmetric unimodal density f can be expressed as a mixture of uniform distributions over symmetric intervals of different lengths; if we define a probability measure on $(0, \infty)$ by $m(dx) = -2xf'(x)dx$, then

$$f(x) = \int_x^\infty \frac{m(da)}{2a}.$$

The result follows.

Thus the density f_n of S_n is symmetric about 0 and unimodal; consequently,

$$\begin{aligned} ||\mathcal{L}(\theta_n S) - \mathcal{L}(\theta_n S')|| &= ||\mathcal{L}(S_n) - \mathcal{L}(S'_n)|| \\ &= 2\mathbb{P}(|S_n| < a/2). \end{aligned}$$

Indeed,

$$\begin{aligned} ||\mathcal{L}(S_n) - \mathcal{L}(S'_n)|| &= \int |f(x - a/2) - f(x + a/2)| \, dx \\ &= 2 \int_0^\infty (f(x - a/2) - f(x + a/2)) \, dx \\ &= 2\, \mathbb{P}(|S_n| < a/2). \end{aligned}$$

By what we have proved in Section 3, it must therefore be the case that

$$\mathbb{P}(|S_n| < a/2) = \mathbb{P}(T > n)$$

for the CM-coupling. To see this, we know already that the optimal coupling Γ_a is described by

$$\begin{aligned} \mathbb{P}(X \in dx, Y = X) &= [f(x) \wedge f(x - a)]dx \quad \text{for } x \in \mathbb{R}; \\ \mathbb{P}(X \in dx, Y = a - X) &= [f(x) - f(x - a)]dx \quad \text{for } x \leq a/2. \end{aligned}$$

So if $h_n(a) \equiv \mathbb{P}[T > n | S_0 = 0, S'_0 = a]$ for $a > 0$, we have the recursion

$$\begin{aligned} h_n(a) &= \int_{-\infty}^{a/2} \{f(x) - f(x - a)\} h_{n-1}(a - 2x) dx \\ &= \int_0^\infty h_{n-1}(t) \{f(\frac{a-t}{2}) - f(\frac{a+t}{2})\} \frac{dt}{2}, \end{aligned}$$

and if we extend the definition of h_n to negative reals by $h_n(-y) = -h_n(y)$ for all $y > 0$, we have more simply for all real a

$$h_n(a) = \int_{-\infty}^{\infty} h_{n-1}(t) f(\frac{a-t}{2}) \frac{dt}{2} \, .$$

Thus if $g(t) \equiv \frac{1}{2} f(\frac{1}{2}t)$, we have the convolution equation

$$h_n = h_{n-1} * g, \qquad n \geq 1,$$

which is solved by

$$h_n(a) = (g^{*n} * h_0)(a) = \mathbb{P}[2S_n < a] - \mathbb{P}[2S_n > a] = \mathbb{P}[|S_n| \leq a/2],$$

as we were expecting.

As examples of this, we could take the symmetric stable random walks. We may think of these as symmetric stable Lévy processes viewed at integer times (see Chapter VIII of Bertoin [1] for more on stable processes).

If the step distribution is symmetric and stable with index $\alpha \in (0, 2]$, we shall have

$$\mathbb{P}[T > n] = \mathbb{P}[|S_n| \leq a/2] = \mathbb{P}[|S_1| \leq \tfrac{1}{2} a n^{-1/\alpha}].$$

For large n, this behaves like $n^{-1/\alpha}$, so for any $\alpha < 1$, the coupling time would have a finite first moment! Note that for $\alpha = 2$, the Gaussian case, we obtain the (typical) rate $O(n^{-1/2})$ of convergence for distributions with finite second moments. To see why this is typical, observe that the rate $O(n^{-1/2})$ can always be achieved for any step distribution (see, for example, Lindvall & Rogers [5]), and we have the trivial bound

$$\begin{aligned} ||\mathcal{L}(\theta_n S) - \mathcal{L}(\theta_n S')|| &\geq |\mathbb{P}(S_n \leq a/2) - \mathbb{P}(S'_n \leq a/2)| \\ &= \mathbb{P}(|S_n| < a/2). \end{aligned}$$

This is $O(n^{-1/2})$ by a local version of the Central Limit Theorem provided the distribution is non-lattice; see, for example, Theorem 2.5.4 in Durrett [3].

REFERENCES

[1] BERTOIN, J.. *Lévy Processes*, Cambridge Univ. Press, Cambridge, 1996.

[2] CUESTA-ALBERTOS, J. A., MATRAN, C., RACHEV, S. T. & RÜSCHENDORF, L.. Mass transportation problems in probability theory, *Math. Scientist,* **21**, 34–72, 1996.

[3] DURRETT, R.. *Probability: Theory and Examples*, Wadsworth & Brooks/Cole, Belmont, Ca., 1991.

[4] GOLDSTEIN, S.. Maximal coupling, *Z. Wahrscheinlichkeitsth. verw. Gebiete* **46**, 193–204, 1979.

[5] LINDVALL, T.. *Lectures on the Coupling Method*, J. Wiley & Sons, New York, 1992.

[6] LINDVALL, T., & ROGERS, L. C. G.. Coupling of multidimensional diffusions by reflection, *Ann. Probability,* **14**, 860–872, 1986.

[7] RACHEV, S. T.. The Monge-Kantorovich mass transference problem and its stochastic applications, *Th. Prob. Applications,* **29**, 647–676, 1985.

[8] THORISSON, H.. On maximal and distributional coupling, *Ann. Probability,* **14**, 873–876, 1986.

School of Mathematical Sciences
University of Bath
Bath BA2 7AY
Great Britain