Skorokhod embeddings for two-sided Markov chains

Peter Mörters* and István Redl
Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, England
E–mail: maspm@bath.ac.uk and ir250@bath.ac.uk

Abstract: Let $(X_n : n \in \mathbb{Z})$ be a two-sided recurrent Markov chain with fixed initial state X_0 and let ν be a probability measure on its state space. We give a necessary and sufficient criterion for the existence of a non-randomized time T such that $(X_{T+n} : n \in \mathbb{Z})$ has the law of the same Markov chain with initial distribution ν. In the case when our criterion is satisfied we give an explicit solution, which is also a stopping time, and study its moment properties. We show that this solution minimizes the expectation of $\psi(T)$ in the class of all non-negative solutions, simultaneously for all non-negative concave functions ψ.

MSc classification (2000): Primary 60J10 Secondary 60G40, 05C70.
Keywords: Skorokhod embedding, stopping time, Markov chain, random walk, extra head scheme, unbiased shift, random allocation, stable matching, optimal transport.

1. Introduction and statement of main results

Let $\mathcal{S}$ be a finite or countable state space and $p = (p_{ij} : i, j \in \mathcal{S})$ an irreducible and recurrent transition matrix. Then there exists a stationary measure $(m_i : i \in \mathcal{S})$ with positive weights, which is finite in the positive recurrent case, and infinite otherwise. The two-sided stationary Markov chain $X = (X_n : n \in \mathbb{Z})$ with initial measure $(m_i : i \in \mathcal{S})$ and transition matrix p is characterized by

- $\mathbb{P}(X_n = i) = m_i$ for all $n \in \mathbb{Z}$, $i \in \mathcal{S}$;
- $\mathbb{P}(X_n = j \,|\, X_{n-1}, X_{n-2}, \dots) = p_{X_{n-1}j}$ for all $n \in \mathbb{Z}$, $i, j \in \mathcal{S}$.

This chain always exists, if we allow $\mathbb{P}$ to be a σ-finite measure. For the simplest construction, let $(X_n : n \geq 0)$ be the chain with initial measure $(m_i : i \in \mathcal{S})$ and transition matrix p, and $(X_{-n} : n \geq 0)$ be the chain with given initial state X_0 and dual transition probabilities given by $p^*_{ij} = (m_j/m_i)p_{ji}$.

By conditioning the stationary chain X on the event $\{X_0 = i\}$, we define the two-sided Markov chain with transition matrix p with fixed initial state $X_0 = i$. Its law, denoted by $\mathbb{P}_i$, does not depend on the choice of $(m_i : i \in \mathcal{S})$ and is always a probability law. Note that we can equivalently define this chain, or indeed the two-sided Markov chain with transition matrix p and arbitrary initial distribution ν, by picking X_0 according to ν and letting the forward and backward chains $(X_n : n \geq 0)$, resp. $(X_{-n} : n \geq 0)$, evolve as in the case of the stationary chain.

A natural version of the *Skorokhod embedding problem* in this context asks, given the two-sided Markov chain $(X_n : n \in \mathbb{Z})$ with transition matrix p and initial state $X_0 = i$ and a probability measure ν on the state space $\mathcal{S}$, whether there exists a random time T such that $(X_{n+T} : n \in \mathbb{Z})$ is a two-sided Markov chain with transition matrix p such that X_T has law ν. If this is the case we say that T is an *embedding* of the *target distribution* ν. Our interest here is mainly in times T which are *non-randomized*, which means that T is a measurable function of the sample chain X. The random times T are often stopping times, but this is not a necessary requirement.

*Communicating author.

Finding embeddings of two-sided Markov chains is a subtle problem, because even for stopping times T the shifted process $T^{-1}X := (X_{n+T} \colon n \in \mathbb{Z})$ often will *not* be a two-sided Markov chain. For example, take a simple symmetric random walk on the integers, started in $X_0 = 0$, and let T be the first positive hitting time of the integer $a > 0$. Then T embeds the Dirac measure δ_a, but the increment $T^{-1}X_0 - T^{-1}X_{-1}$ always takes the value $+1$, hence $T^{-1}X$ is not a two-sided simple random walk. A similar argument shows that even shifting the simple random walk by a nonzero fixed time does not preserve the property of being a simple random walk with given distribution of the state at time zero.

The first main result of this paper gives a necessary and sufficient condition on the initial state, the target measure and the stationary distribution for the existence of a Skorokhod embedding for an arbitrary two-sided Markov chain.

Theorem 1. *Let X be a two-sided irreducible and recurrent Markov chain with transition matrix p and initial state $X_0 = i$. Take $\nu = (\nu_j \colon j \in \mathcal{S})$ to be any probability measure on $\mathcal{S}$. Then the following statements are equivalent.*

(a) *There exist a non-randomized random time T such that $(X_{n+T} \colon n \in \mathbb{Z})$ is a Markov chain with transition matrix p and X_T has law ν.*

(b) *The stationary measure $(m_j \colon j \in \mathcal{S})$ satisfies $\frac{m_i}{m_j}\, \nu_j \in \mathbb{Z}$ for all $j \in \mathcal{S}$.*

If the random time T in (a) exists it can always be taken to be a stopping time.

Example 1.1 (Embedding measures with mass in the initial state) Assume that the target measure ν charges the initial state $i \in \mathcal{S}$ of the Markov chain, i.e. $\nu_i > 0$. Choosing $i = j$ in (b) shows that a non-randomized random time T with the properties of (a) can exist only if $\nu = \delta_i$. In this case a natural family of embeddings can be constructed using the concept of point stationarity, see for example [19], as follows: Let $r \in \mathbb{N}$ and let T_r be the the time of the rth visit of state i after time zero. Then it is easy to check, and follows from [14, Theorem 6.3], that the process $T_r^{-1}X$ is a Markov chain with transition matrix p and $X_{T_r} = i$. ■

Example 1.2 (Extra head problem) Take a doubly-infinite sequence of tosses of a (possibly biased) coin, or more precisely let $X = (X_n \colon n \in \mathbb{Z})$ be i.i.d. random variables with distribution $\mathbb{P}(X_n = \texttt{head}) = p$, $\mathbb{P}(X_n = \texttt{tail}) = 1 - p$, for some $p \in (0, 1)$. Our aim is to find, without using any randomness generated in a way different from looking at coins in the sequence, a coin showing head in this sequence in such a way that the two semi-infinite sequences of coins to the left and to the right of this coin remain independent i.i.d. sequences of coins with the same bias. This is known as *extra head problem* and was investigated and fully answered by Liggett [15] and Holroyd and Peres [12]. To relate this to our setup, we can assume that $X_0 = \texttt{tail}$, as otherwise the coin at the origin is the extra head. Then the extra head problem becomes the Skorokhod embedding problem for X with initial state $X_0 = \texttt{tail}$ and target measure $\nu = \delta_{\texttt{head}}$. Theorem 1 shows (as proved by Holroyd and Peres before) that the extra head problem has a solution if and only if $(1 - p)/p \in \mathbb{Z}$, i.e. if and only if p is the inverse of an integer. Moreover, Liggett [15] gives an explicit solution of the extra head problem which we generalize to our setup in Theorem 2 below. ■

Example 1.3 (Inverse extra head problem) If in the setup of Example 1.2 the state of the coin at the origin has been revealed, we ask whether it is possible to shift the sequence in such a way that this information is lost, i.e. the shifted sequence is an i.i.d. sequence of coins with the original bias. This means that we wish to embed the invariant distribution $\nu = m$ given by $m_{\texttt{head}} = p, m_{\texttt{tail}} = 1 - p$. Theorem 1 shows that this is impossible. ■

Example 1.4 (Extra head problem with a finite pattern) In the setup of Example 1.2 we now ask to find a particular *finite pattern* of successive outcomes, such that the coins to its left and right remain an i.i.d. sequence of coins with the same bias. Looking, for example, for the pattern `head/tail` we would first reveal the coin at the origin, and then if this shows `head` its right neighbour, and if this shows `tail` its left neighbour. The underlying Markov chain has the state space $\{\texttt{tail/tail}, \texttt{tail/head}, \texttt{head/tail}, \texttt{head/head}\}$, the transition matrix

$$\begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \end{pmatrix},$$

and invariant measure $((1-p)^2,\, p(1-p),\, p(1-p),\, p^2)$. Our theorem shows that, if we initially reveal `tail/tail` then we need $1/p$ to be an integer, and if we reveal `head/head` then we need $1/(1-p)$ to be an integer. Hence we can only embed `head/head` if $p = \frac{1}{2}$. More generally, the problem can be solved for patterns that are repetitions of the single symbol `head` if and only if $1/p$ is an integer, for patterns that are repetitions of the single symbol `tail` if and only if $1/(1-p)$ is an integer, and for patterns containing both symbols `tail` and `head` if and only if $p = \frac{1}{2}$. ■

Example 1.5 (Simple random walk) Let X be a two-sided simple symmetric random walk on the integers, with $X_0 = i$ for some $i \in \mathbb{Z}$. In this case the invariant measure is $m_i = 1$ for all $i \in \mathbb{Z}$, hence Theorem 1 shows that the target measures that can be embedded are precisely the Dirac measures δ_j, $j \in \mathcal{S}$. The same result holds for the simple symmetric random walk on the square lattice $\mathbb{Z}^2$. ■

The proof of Theorem 1 extends the ideas developed by Liggett [15] and Holroyd and Peres [12] for the extra head problem to the more general Markov chain setup. In particular, under the additional assumption that the target measure does not charge the initial state, we are able to generalize Liggett's construction of an elegant explicit solution, in analogy to the Brownian motion case studied in Last et al. [13]. Recall that the case when the target measure charges the initial state was already discussed in Example 1.1. To describe this solution we define the local time L^j spent by X at state $j \in \mathcal{S}$ to be the normalized counting measure given by

$$L^j(A) := \frac{1}{m_j}\, \#\{n \in A \colon X_n = j\} \qquad \text{for any } A \subset \mathbb{Z}.$$

Theorem 2. *Let X be a two-sided irreducible and recurrent Markov chain with $X_0 = i$ and further assume that the target measure ν satisfies $\nu_i = 0$ and the conditions in Theorem 1(b). Then*

$$T_* := \min\Big\{ n \geq 0 \colon L^i([0,n]) \leq \sum_{j \in \mathcal{S}} \nu_j\, L^j([0,n]) \Big\} \tag{1.1}$$

is a finite, non-randomized stopping time satisfying the conditions of Theorem 1(a).

Example 1.6 We take a stationary three state Markov chain with transition probabilities given by $p_{12} = p_{32} = 1$ and $p_{21} = 1-p$ and $p_{23} = p$. If $1/p$ is an integer we can shift the chain so that it starts in the third state and the chain property is preserved, as follows: Uncover the state at the origin. If it is the third state we are done; if it is the second state we move along the chain until the number of visits to the third state is at least p times the number of visits to the second state; if it is the first state we move until the number of visits to the third state is at least $\frac{p}{1-p}$ times the number of visits to the first state. Note that if the state of the origin is the first state it is *not* a solution to wait one time step, whence you are in the second state, and then apply the strategy for start in the second state as this creates a bias in the backward chain. ■

Skorokhod embedding problems usually concern embedding times with finite expectation. However in the extra head problem it is not possible to achieve finite expectation of the random time T. In fact Liggett [15] shows that in this case always $E\sqrt{T} = \infty$, see also Holroyd and Liggett [10]. For the simple random walk on the integers we expect in analogy to the Brownian motion case studied by Last et al. [13] that always $E\sqrt[4]{T} = \infty$. Our aim here is to understand the general picture.

To this end we now recall the notion of asymptotic Green's function of the Markov chain. Given states $i, j \in \mathcal{S}$ we first define the normalized truncated Green's function by

$$a_{ij}(n) = \mathbb{E}_i L^j([0,n]) = \frac{1}{m_j}\,\mathbb{E}_i\Big[\sum_{k=0}^{n} \mathbb{1}\{X_k = j\}\Big],$$

that is $a_{ij}(n)$ gives the normalized expected number of visits to state j between time 0 and time n, by the Markov chain with initial state $X_0 = i$. By Orey's ergodic theorem, see, e.g., Chen [5], for any states $i, j, k, l \in \mathcal{S}$, the functions a_{ij} and a_{kl} are asymptotically equivalent in the sense that

$$\lim_{n\to\infty} \frac{a_{ij}(n)}{a_{kl}(n)} = 1.$$

We then define the *asymptotic Green's function* $a(n)$ as the equivalence class of the truncated Green's functions under asymptotic equivalence. Observe that finiteness of moments is a class property, i.e. expressions of the form $E[a(Y)] < \infty$, where a is an equivalence class and Y an integer-valued random variable, are meaningful.

Theorem 3. *Let X be a two-sided irreducible and recurrent Markov chain with $X_0 = i$ and ν be any target measure different from the Dirac measure δ_i. If T_* is the stopping time defined in* (1.1)*, then*

(i) $\mathbb{E}_i\big[a(T_*)^{1/2}\big] = \infty$.

If additionally ν has finite support, then

(ii) $\mathbb{E}_i\big[a(T_*)^{\beta}\big] < \infty$ *for all* $0 \le \beta < \frac{1}{2}$.

As $a(n)$ cannot grow faster than n, our solutions T_* always have 'bad' moment properties as even for the nicest Markov chain T_* can never have finite square root moments. However, our next theorem shows that no other solution of the embedding problem has better moment properties than T_*.

In fact, it turns out that T_* has a strong optimality property, as it simultaneously minimizes all concave moments of non-negative solutions of the embedding problem. This striking result is new even for the case of the extra head problem and therefore, in our opinion, constitutes the most interesting contribution in this paper.

Theorem 4. *Let X be a two-sided irreducible and recurrent Markov chain with $X_0 = i$ and ν be a target measure satisfying the conditions in Theorem 2. If T_* is the solution of the Skorokhod embedding problem constructed in* (1.1) *and T any other non-negative (possibly randomized) solution, then*

$$\mathbb{E}_i\big[\psi(T_*)\big] \le \mathbb{E}_i^{\oplus}\big[\psi(T)\big],$$

for any non-negative concave function ψ defined on the non-negative integers, where the expectation on the right is with respect to the chain as well as any possible extra randomness used to define T.

Theorem 4 is inspired by exciting recent developments connecting the classical Skorokhod embeddings for Brownian motion with optimal transport problems. In a recent paper, Beiglböck, Cox and Huesmann [4] exploit this connection to characterize certain solutions to the Skorokhod embedding problem by a geometric property. In a similar vein, our solution T_* is characterized by a geometric property, the 'non-crossing' condition, which yields the optimality. See also our concluding remarks in Section 6 for possible extensions of this result.

Example 1.7 Suppose the underlying Markov chain is *positive recurrent*. Then the asymptotic Green's function satisfies $a(n) \sim n$. Therefore all non-negative solutions T of the Skorokhod embedding problem satisfy $\mathbb{E}_i[\sqrt{T}] = \infty$, while the solution constructed in Theorem 2 satisfy $\mathbb{E}_i[T_*^\beta] < \infty$ for all $0 \leq \beta < 1/2$. This applies in particular to Examples 1.2 and 1.4. ■

Example 1.8 The situation is much more diverse for *null-recurrent* chains. Looking at Example 1.5, for a two-sided simple symmetric random walk on the integers we have $a(n) \sim \sqrt{n}$. Hence the solution T_* constructed in Theorem 2 satisfies $\mathbb{E}_i[T_*^\alpha] < \infty$ for all $0 \leq \alpha < 1/4$, while any non-negative solution has infinite $1/4$ moment. This is similar to the case of Brownian motion on the line, which is discussed in [13], although in that paper other than here the discussion is restricted to solutions which are non-randomized stopping times. In contrast to this, for simple symmetric random walk on the square lattice $\mathbb{Z}^2$ we have $a(n) \sim \log n$, and therefore $\mathbb{E}_i[\sqrt{\log T}]$ is infinite for any non-negative solution T, while the solution T_* constructed in Theorem 2 satisfies $\mathbb{E}_i[(\log T_*)^\alpha] < \infty$, for all $0 \leq \alpha < 1/2$. ■

2. Relating embedding and allocation problems

In this section we relate our embedding problem to an equivalent allocation problem. The section specializes some results from Last and Thorisson [14] which are themselves based on ideas from [12]. We give complete proofs of the known facts in order to keep this paper self-contained. Generalizing from [13] we call a random time T an *unbiased shift* of the Markov chain X if the shifted process $T^{-1}X$ is a two-sided Markov chain with the same transition matrix as X. Note that this definition allows T to be randomized, i.e. it does not have to be a function of the sample chain X alone.

Let $\Omega = \{(\omega_i)_{i\in\mathbb{Z}} \colon \omega_i \in \mathcal{S}\}$ be the set of trajectories of X. A *transport rule* is a measurable function $\theta \colon \Omega \times \mathbb{Z} \times \mathbb{Z} \to [0,1]$ satisfying

$$\sum_{y\in\mathbb{Z}} \theta_\omega(x,y) = 1 \qquad \text{for all } x \in \mathbb{Z} \text{ and } \mathbb{P}\text{-almost every } \omega.$$

Note that we write the dependence on the trajectory ω by a subindex, which we drop from the notation whenever convenient. Transport rules are interpreted as distributing mass from x to $\mathbb{Z}$ in such a way that the site y gets a proportion $\theta(x,y)$ of the mass. For sets $A, B \subset \mathbb{Z}$ we define

$$\theta_\omega(A,B) := \sum_{x\in A, y\in B} \theta_\omega(x,y).$$

A transport rule θ is called *translation invariant* if

$$\theta_{z\omega}(x+z, y+z) = \theta_\omega(x,y),$$

for all $\omega \in \Omega$ and $x, y, z \in \mathbb{Z}$, where $z\omega$, defined by $z\omega_n = \omega_{n-z}$ for any $n \in \mathbb{Z}$, is the trajectory shifted by $-z$. A transport rule *balances* the random measures ξ and ζ on $\mathbb{Z}$ if

$$\sum_{z\in\mathbb{Z}} \theta_\omega(z,A)\xi(z) = \zeta(A), \tag{2.1}$$

for any $A \subset \mathbb{Z}$ and $\mathbb{P}$-almost all ω. Given a two-sided Markov chain X as before recall the definition of the local times L^i, and given a probability measure $\nu = (\nu_i \colon i \in \mathcal{S})$ we further define

$$L^\nu = \sum_{i \in \mathcal{S}} \nu_i \, L^i.$$

Proposition 2.1. *Assume that there is a measurable family of probability measures $(\mathbb{Q}_\omega \colon \omega \in \Omega)$ on some measurable space Ω' and $T \colon \Omega \times \Omega' \to \mathbb{Z}$ is measurable. The random time T and a translation invariant transport rule θ are associated if*

$$\mathbb{Q}_\omega\big(\omega' \in \Omega' \colon T(\omega, \omega') = t\big) = \theta_\omega(0, t) \qquad \text{for all } t \in \mathbb{Z} \text{ and } \mathbb{P}\text{-almost all } \omega \in \Omega. \tag{2.2}$$

For any probability measure $\mu = (\mu_i \colon i \in \mathcal{S})$ we define the probability measure $\mathbb{P}^\oplus_\mu$ on $\Omega \times \Omega'$ by

$$\mathbb{P}^\oplus_\mu(d\omega \, d\omega') = \sum_{i \in \mathcal{S}} \mu_i \, \mathbb{P}_i(d\omega) \, \mathbb{Q}_\omega(d\omega'). \tag{2.3}$$

Then, if μ, ν is any pair of probability measures on $\mathcal{S}$ and the random time T and translation invariant transport rule θ are associated, the following statements are equivalent.

- *(a) Under $\mathbb{P}^\oplus_\mu$ the random time T is an unbiased shift of X and X_T has law ν.*
- *(b) The transport rule θ balances L^μ and L^ν $\mathbb{P}$-almost everywhere.*

Note that in the last proposition unbiased shifts need not be non-randomized. The transport rules associated to non-randomized shifts are the *allocation rules.* These are given by a measurable map $\tau \colon \Omega \times \mathbb{Z} \to \mathbb{Z}$ such that $\theta_\omega(x, y) = 1$ if $\tau_\omega(x) = y$ and zero otherwise.

Proposition 2.2. *If the random time T in Proposition 2.1 is non-randomized, then there is an associated transport rule θ, which is an allocation rule. Conversely if θ in Proposition 2.1 is an allocation rule, then there exists an associated non-randomized random time T.*

We give proofs of the propositions for completeness. For a transport rule θ we define

$$J_\mu(\omega) := \sum_{k \in \mathbb{Z}} \theta_\omega(k, 0) \, L^\mu(k), \tag{2.4}$$

which is interpreted as the total mass received by the origin. We recall the following simple fact, see [12] for a more general version.

Lemma 2.3. *Let $m \colon \mathbb{Z} \times \mathbb{Z} \to [0, \infty]$ be such that $m(x + z, y + z) = m(x, y)$ for all $x, y, z \in \mathbb{Z}$. Then*

$$\sum_{y \in \mathbb{Z}} m(x, y) = \sum_{y \in \mathbb{Z}} m(y, x).$$

The following calculation is at the core of the proof.

Lemma 2.4. *Suppose that T and θ are related by (2.2). Then, for any measurable function $f \colon \Omega \to [0, \infty]$, we have*

$$\mathbb{E}^\oplus_\mu \big[f(T^{-1}X)\big] = \mathbb{E}\big[J_\mu(X) f(X)\big],$$

where $\mathbb{E}^\oplus_\mu$ is the expectation with respect to $\mathbb{P}^\oplus_\mu$ defined in (2.3).

Proof of Lemma 2.4. Writing $\mathbb{P}_\mu = \sum_{i \in \mathbb{Z}} \mu_i \mathbb{P}_i$ we get

$$\begin{aligned} \mathbb{E}^\oplus_\mu \big[f(T^{-1}X)\big] &= \int d\mathbb{P}_\mu(\omega) \int f\big(T(\omega, \omega')^{-1} X(\omega)\big) \, \mathbb{Q}_\omega(d\omega') \\ &= \int d\mathbb{P}_\mu(\omega) \sum_{t \in \mathbb{Z}} \mathbb{Q}_\omega(T = t) f(t^{-1} X(\omega)). \end{aligned}$$

Using relation (2.2) and the definition of $\mathbb{P}_\mu$ we continue with

$$= \sum_{i\in\mathbb{Z}} \mu_i \int d\mathbb{P}_i(\omega) \sum_{t\in\mathbb{Z}} \theta_\omega(0,t) f(t^{-1}X(\omega))$$
$$= \sum_{i\in\mathbb{Z}} \mu_i \int d\mathbb{P}(\omega) \sum_{t\in\mathbb{Z}} \theta_\omega(0,t) L^i(0) f(t^{-1}X(\omega)),$$

as $L^i(0) = \frac{1}{m_i}$ and $L^j(0) = 0$ under $\mathbb{P}_i$ for $j \neq i$. Applying Lemma 2.3 gives

$$= \sum_{i\in\mathbb{Z}} \mu_i \int d\mathbb{P}(\omega) \sum_{t\in\mathbb{Z}} \theta_\omega(t,0) L^i(t) f(X(\omega))$$
$$= \int d\mathbb{P}(\omega) \sum_{t\in\mathbb{Z}} \theta_\omega(t,0) L^\mu(t) f(X(\omega))$$
$$= \mathbb{E}\big[J_\mu(X) f(X)\big],$$

using first the definition of L^μ and second the definition of $J_\mu(X)$. □

Proof of Proposition 2.1. First assume that θ is a translation invariant transport rule. Then, for any non-negative measurable f, by Lemma 2.4, we have

$$\mathbb{E}^{\oplus}_\mu\left[f(T^{-1}X)\right] = \mathbb{E}\big[J_\mu(X) f(X)\big] = \mathbb{E}\big[\sum_{k\in\mathbb{Z}} \theta_\omega(k,0)\, L^\mu(k) f(X)\big]. \tag{2.5}$$

If θ balances L^μ and L^ν this equals

$$\mathbb{E}\big[L^\nu(0) f(X)\big] = \sum_{j\in\mathbb{Z}} \nu_j\, \mathbb{E}\big[L^j(0) f(X)\big] = \sum_{j\in\mathbb{Z}} \nu_j\, \mathbb{E}_j\big[f(X)\big] = \mathbb{E}_\nu\big[f(X)\big].$$

Hence under $\mathbb{P}^{\oplus}_\mu$ the random variable $T^{-1}X$ has the law of X under $\mathbb{P}_\nu$. In other words T is an unbiased shift and X_T has distribution ν.

Conversely, assume that T is an unbiased shift and X_T has distribution ν. Hence $\mathbb{E}^{\oplus}_\mu[f(T^{-1}X)] = \mathbb{E}_\nu[f(X)] = \mathbb{E}[L^\nu(0) f(X)]$. Plugging this into (2.5) gives

$$\mathbb{E}\big[\sum_{k\in\mathbb{Z}} \theta_\omega(k,0)\, L^\mu(k) f(X)\big] = \mathbb{E}\big[L^\nu(0) f(X)\big].$$

As f was arbitrary we get $\sum_{k\in\mathbb{Z}} \theta_\omega(k,0)\, L^\mu_\omega(k) = L^\nu_\omega(0)$ for $\mathbb{P}$-almost all ω, where we emphasise the dependence of the measures L^μ and L^ν on the trajectories by a subscript. As θ is translation invariant we get, substituting $m := k - \ell$,

$$\sum_{k\in\mathbb{Z}} \theta_\omega(k,A)\, L^\mu_\omega(k) = \sum_{k\in\mathbb{Z}} \sum_{\ell\in A} \theta_\omega(k,\ell)\, L^\mu_\omega(k) = \sum_{\ell\in A} \sum_{m\in\mathbb{Z}} \theta_{-\ell\omega}(m,0)\, L^\mu_{-\ell\omega}(m)$$
$$= \sum_{\ell\in A} L^\nu_{-\ell\omega}(0) = \sum_{\ell\in A} L^\nu_\omega(\ell) = L^\nu_\omega(A),$$

for every $A \subset \mathbb{Z}$ and $\mathbb{P}$-almost every ω. □

Proof of Proposition 2.2. Suppose $T = T(\omega)$ is non-randomized. Define $\tau_\omega\colon \mathbb{Z} \to \mathbb{Z}$ by $\tau_\omega(k) = T(-k\omega) + k$ and let $\theta_\omega(x,y) = 1$ if $\tau_\omega(x) = y$ and zero otherwise. Then θ is a translation invariant allocation rule. Moreover, $\mathbb{Q}_\omega(T = t) = \mathbb{1}\{t = T(\omega)\} = \mathbb{1}\{t = \tau_\omega(0)\} = \theta_\omega(0,t)$, hence T and θ are associated. Conversely, if θ is a translation invariant allocation rule given by $\tau\colon \Omega \times \mathbb{Z} \to \mathbb{Z}$ define a non-randomized time T by $T = \tau_\omega(0)$. As before, $\mathbb{Q}_\omega(T = t) = \mathbb{1}\{t = T(\omega)\} = \mathbb{1}\{t = \tau_\omega(0)\} = \theta_\omega(0,t)$, and hence T and θ are associated. □

3. Existence of allocation rules: Proof of Theorems 1 and 2

In the light of the previous section our Theorems 1 and 2 can be formulated and proved as equivalent statements about allocation rules. We start with the result on non-existence of non-randomized unbiased shifts, which is implicit in Theorem 1.

Suppose that statement (a) in Theorem 1 holds and for the Markov chain X with $X_0 = i$ there exists a non-randomized unbiased shift T such that X_T has law ν. Then by Proposition 2.2 there exists a translation-invariant allocation rule τ associated with T and by Proposition 2.1 this rule balances the measures L^i and L^ν. Recall that L^i is the measure on $\mathbb{Z}$ which has masses of fixed size $1/m_i$ at the times when the stationary chain X visits state i. By the balancing property (2.1) for allocation rules, all masses of L^ν must have sizes which are integer multiples of $1/m_i$. As these masses are ν_j/m_j we get that $\frac{m_i}{m_j}\nu_j$ must be integers for all $j \in \mathcal{S}$, which is statement (b).

The remainder of this section is devoted to the proof of *existence* of non-randomized unbiased shifts of the Markov chain X with $X_0 = i$, embedding ν under the assumption of Theorem 1 (b). By Example 1.1 we may additionally assume that for the initial state i of the Markov chain we have $\nu_i = 0$. Our claim is that the stopping time T_* defined in Theorem 2 is an unbiased shift with the required properties. The next proposition shows that an associated allocation rule balances the measures L^i and L^ν which, once accomplished, implies Theorem 2 and completes the proof of Theorem 1.

Proposition 3.1. *Under the assumptions set out above, the following holds.*

(a) *The mapping $\tau\colon \Omega \times \mathbb{Z} \to \mathbb{Z}$ defined by*

$$\tau_\omega(k) = \min\big\{n \geq k\colon L^i_\omega([k,n]) \leq L^\nu_\omega([k,n])\big\}$$

is a translation-invariant allocation rule associated with the T_* defined in (1.1).

(b) *For $\mathbb{P}$-almost every ω and all $A \subset \mathbb{Z}$ we have*

$$\sum_{k\in\mathbb{Z}} \mathbb{1}\{\tau_\omega(k) \in A\}\, L^i_\omega(k) = L^\nu_\omega(A), \tag{3.1}$$

in other words the allocation rule balances L^i and L^ν.

The proof of the proposition is similar to that of [13, Theorem 5.1] in the diffuse case. We prepare it with two lemmas. The first lemma is a pathwise statement which holds for every fixed trajectory ω satisfying the stated assumption.

Lemma 3.2. *Suppose $b \in \mathbb{Z}$ is such that $X_b = j$ for some $j \in \mathcal{S}$ with $\nu_j > 0$, and $a \in \mathbb{Z}$ is given by*

$$a := \max\big\{k < b\colon L^i([k,b]) \geq L^\nu([k,b])\big\}.$$

Then

$$\sum_{k\in[a,b]} \mathbb{1}\{\tau(k) \in A\}\, L^i(k) = L^\nu(A), \tag{3.2}$$

holds for any $A \subset [a,b]$.

Proof. We define the function $\Delta f\colon \mathbb{Z} \to [0,\infty)$ by

$$\Delta f(k) := L^i(k) - L^\nu(k) = \begin{cases} \frac{1}{m_i} & \text{if} \quad X_k = i, \\ -\frac{\nu_j}{m_j} & \text{if} \quad X_k = j \neq i. \end{cases}$$

Recall that by our assumption $\frac{\nu_j}{m_j}$ is an integer multiple of $\frac{1}{m_i}$. Hence, denoting

$$f_u^v := \sum_{n=u}^{v} \Delta f(n) \quad \text{ for all } u, v \in \mathbb{Z} \text{ and } u \le v,$$

we have $a = \max\{k < b \colon f_k^b = 0\}$ and hence $f_a^b = 0$.

By the additivity of both sides of (3.2) it suffices to prove

$$\sum_{k\in[a,b]} \mathbb{1}\{\tau(k) = z\}\, L^i(k) = L^\nu(z) \quad \text{for all sites } z \in [a,b]. \tag{3.3}$$

Fix $z \in [a,b]$ and let $j = X_z$. Observe that $\tau(k) = z$ if and only if $f_k^z \le 0$ but $f_k^\ell > 0$ for all $k \le \ell < z$. Hence we may assume $\nu_j > 0$ as otherwise both sides of (3.3) are zero. We also have that $f_a^z > 0$ if $z < b$. Indeed, suppose that $f_a^z \le 0$. Then $f_{z+1}^b = f_a^b - f_a^z \ge 0$ contradicting the choice of a.

As $f_a^z \ge 0, f_z^z = -\frac{\nu_j}{m_j} < 0$ and ν_j/m_j is an integer multiple of $1/m_i$ we find a $k_1 \ge a$ with $f_{k_1}^z = 0$ and $f_j^z < 0$ for all $k_1 < j \le z$. Similarly, we find $k_1 < k_2 < \cdots < k_N$ where $N := (\frac{m_i}{m_j})\, \nu_j$ such that

$$f_{k_n}^z = \tfrac{1-n}{m_i} \text{ and } f_j^z < \tfrac{1-n}{m_i} \text{ for all } k_n < j \le z.$$

As $\tau(k) = \min\{n \ge k \colon f_k^n \le 0\}$ we infer that $\tau(k_n) = z$ for all $n \in \{1, \ldots, N\}$ and there are no other values k with $\tau(k) = z$. Each of these values contributes a summand $\frac{1}{m_i}$ to the left hand side in (3.3). Therefore this side equals $\frac{N}{m_i} = \frac{\nu_j}{m_j}$, as does the right hand side. This completes the proof. □

The second lemma is probabilistic and ensures in particular that the mapping τ described in Proposition 3.1 (a) is well defined.

Lemma 3.3. *For $\mathbb{P}$-almost every ω the following two events hold*

- *(E1) for all k with $X_k = i$ we have $\tau(k) < \infty$;*
- *(E2) for all b such that $X_b = j$ for some $j \in \mathcal{S}$ with $\nu_j > 0$ there exists $a < b$ such that $X_a = i$ and $L^i([a,b]) = L^\nu([a,b])$.*

Proof. To show this we use an argument from [12], see Theorem 17 and the following remark. We formulate the negation of the two events. The complement of (*E1*) is the event that there exists k such that $X_k = i$ and $L^i([k,\ell]) > L^\nu([k,\ell])$, for all $\ell > k$. The complement of (*E2*) is that there exists b such that $X_b = j$ for some $j \in \mathcal{S}$ with $\nu_j > 0$ and $L^i([a,b]) < L^\nu([a,b])$, for all $a < b$ with $X_a = i$. We first show that, for $\mathbb{P}$-almost every ω, both complements cannot occur simultaneously.

Indeed, for a fixed ω, it is clear that there cannot be k and b as above such that $k < b$. Assume for contradiction that the set of trajectories ω for which there exist $k > b$ as above has positive probability. On this event the minimum over all k with $\tau(k) = \infty$ for all $\ell > k$ is finite, we denote it by K. By translation invariance $\mathbb{P}(K = 0) > 0$ from which we infer by conditioning on the event $\{X_0 = i\}$ that $\mathbb{P}_i(K = 0) > 0$. If $(T_n \colon n \in \mathbb{N})$ is the collection of return times to state i, by the invariance described in Example 1.1 we have $\mathbb{P}_i(K = T_n) = \mathbb{P}_i(K = 0) > 0$ for all $n \in \mathbb{N}$ contradicting the finiteness of $\mathbb{P}_i$. Therefore we have shown that, for $\mathbb{P}$-almost every ω, either (*E1*) or (*E2*) occurs.

As the last step we show that event (*E1*) cannot occur without event (*E2*). To this end define $m(x,y) = \mathbb{E}[\mathbb{1}\{\tau(x) = y, X_x = i\}]$ and apply Lemma 2.3 to get

$$\mathbb{E}\Big[\sum_{k\in\mathbb{Z}} \mathbb{1}\{\tau(k) = 0, X_k = i\}\Big] = \mathbb{E}\Big[\sum_{k\in\mathbb{Z}} \mathbb{1}\{\tau(0) = k, X_0 = i\}\Big].$$

The left-hand side in this equation equals m_i if and only if $(E2)$ occurs $\mathbb{P}$-almost every ω, and the right-hand side equals m_i if and only if $(E1)$ occurs $\mathbb{P}$-almost every ω. As these two events cannot fail at the same time, both events $(E1)$ and $(E2)$ occur for $\mathbb{P}$-almost every ω. □

Proof of Proposition 3.1. Recall that τ is well-defined and note that translation-invariance of the allocation rule defined in terms of τ follows easily from the fact that $\tau_\omega(k) = \tau_{k\omega}(0) + k$. As $T_*(\omega) = \tau_\omega(0)$ by definition, the allocation rule is associated with T_*. This proves (a).

To prove (b) we note that it suffices to fix $z \in \mathbb{Z}$ and show that for $\mathbb{P}$-almost every ω equation (3.1) holds for $A = \{z\}$. We let $b = \tau(z)$. By Lemma 3.3 for $\mathbb{P}$-almost every ω there exists $a < b$ such that $X_a = i$ and $L^i([a,b]) = L^\nu([a,b])$. Then the interval $[a,b]$ contains z and all k with $\tau(k) = z$. Hence the results follows by application of Lemma 3.2. □

4. Moment properties of T_*: Proof of Theorem 3

The critical exponent $\frac{1}{2}$ occurring in Theorem 3 originates from the behaviour of the first passage time below zero by a mean zero random walk. We summarize the results required for such random walks in the following lemma.

Lemma 4.1. *Let $\xi, \xi_1, \xi_2, \dots$ be independent identically distributed random variables with $E\xi = 0$ taking values in the integers. Define the associated random walk by $S_n = \sum_{i=1}^n \xi_i$ and its first passage time below zero as $N = \min\{n \in \mathbb{N} \colon S_n \leq 0\}$.*

- *(a) If the walk is skip-free to the right, i.e. $P(\xi > 1) = 0$, then $E\big[N^{1/2}\big] = \infty$.*
- *(b) If the walk has finite variance, then there exists $C > 0$ such that $P\big(N > n\big) \sim C\,\dfrac{1}{\sqrt{n}}$.*

Proof. (a) Denote by $N^{(j)}$ the first passage time for the walk given by $S_n^{(j)} = \sum_{i=1}^n \xi_{i+j-1}$. Then

$$E[N \wedge n] = \sum_{j=1}^n P(N \geq j) = \sum_{j=1}^n P(N^{(j)} \geq n-j+1) = E\Big[\sum_{j=1}^n 1\{N^{(j)} \geq n-j+1\}\Big],$$

If $\underline{S}_n$ denotes the minimum of $\{S_0, S_1, \dots, S_n\}$ we have, using that the walk is skip-free to the right,

$$\sum_{j=1}^n 1\{N^{(j)} \geq n-j+1\} = S_n - \underline{S}_n.$$

This implies $E[N \wedge n] \geq E[(S_n)_+]$. By a concentration inequality for arbitrary sums of independent random variables, see [16, Theorem 2.22], there exists a constant $C > 0$ such that, for all $\varepsilon > 0$ and $n \in \mathbb{N}$, we have $P(S_n \in [-\varepsilon\sqrt{n}, \varepsilon\sqrt{n}]) \leq C\,\varepsilon$. Hence, by Markov's inequality, for any $\varepsilon > 0$,

$$E\big[(S_n)_+\big] = \frac{1}{2} E|S_n| \geq \frac{1}{2}\,\varepsilon\sqrt{n}\,P\big(|S_n| > \varepsilon\sqrt{n}\big) \geq \frac{1}{2}\,\varepsilon(1 - C\varepsilon)\,\sqrt{n}.$$

We infer that $\liminf \frac{1}{\sqrt{n}}\,E[N \wedge n] > 0$. But if we we had $E[N^{1/2}] < \infty$ dominated convergence would imply that this limit is zero, which is a contradiction.

(b) This is a classical result of Spitzer [18]. A good proof can be found in [8, Theorem 1a in Section XII.7], see also [8, Section XVIII.5] for a proof that random walks with finite variance satisfy Spitzer's condition. □

4.1 Proof of Theorem 3 (i). We start by proving a variant of the upper half in the Barlow-Yor inequality [2] for Markov chains. This result, usually given in the context of continuous martingales, estimates the moments of the local time at a stopping time, by moments of the stopping time itself.

Lemma 4.2. *For any $0<p<\infty$, there exists a constant C_p such that, for any state $i\in\mathcal{S}$ and any stopping time T,*

$$\mathbb{E}_i[L^i([0,T])^p] \le C_p\,\mathbb{E}_i[a_{ii}(T)^p]. \tag{4.1}$$

The lemma relies on the following classical inequality, we refer to [3, (6.9)] for a proof.

Lemma 4.3 (Good λ inequality). *For every $0<p<1$ there is a constant $C_p>0$ such that, for any pair of non-negative random variables (X,Y) satisfying*

$$P(X>3\lambda, Y<\delta\lambda)\le \delta\,P(X>\lambda) \quad \text{for all } 0<\delta<3^{-p-1} \text{ and } \lambda>0, \tag{4.2}$$

we have

$$E\left[X^p\right]\le C_p\,E\left[Y^p\right].$$

Proof of Lemma 4.2. If we show that (4.2) holds with random variables $X=m_iL^i([0,T])$ and $Y=m_ia_{ii}(T)$ under $\mathbb{P}_i$, the result follows immediately from Lemma 4.3. If $\lambda\le 1$ the left hand side of (4.2) is zero and there is nothing to show. We may therefore assume that $\lambda>1$. Define $m_ia_{ii}^{-1}(x):=\max\{n\colon m_ia_{ii}(n)<x\}$. Let $T_0=0$ and T_k be the time of the kth visit of state i after time zero. Finally assume, without loss of generality, that $\mathbb{P}_i(X>\lambda)>0$. Then,

$$\begin{aligned}\mathbb{P}_i\big(X>3\lambda, Y<\delta\lambda\,\big|\,X>\lambda\big)&=\mathbb{P}_i\big(T_{\lfloor 3\lambda\rfloor+1}\le T, m_ia_{ii}(T)<\delta\lambda\,\big|\,T_{\lfloor\lambda\rfloor+1}\le T\big)\\ &\le \mathbb{P}_i\big(T_{\lfloor 3\lambda\rfloor-\lfloor\lambda\rfloor}\le m_ia_{ii}^{-1}(\delta\lambda)\big)\le \mathbb{P}_i\big(L^i([0,m_ia_{ii}^{-1}(\delta\lambda)])\ge\lfloor 2\lambda\rfloor\big).\end{aligned}$$

By Markov's inequality the last expression above can be bounded by

$$\lfloor 2\lambda\rfloor^{-1}\,\mathbb{E}_i\left[L^i([0,m_ia_{ii}^{-1}(\delta\lambda)])\right]=\lfloor 2\lambda\rfloor^{-1}\,m_ia_{ii}\left(m_ia_{ii}^{-1}(\delta\lambda)\right)\le\delta\frac{\lambda}{\lfloor 2\lambda\rfloor},$$

which is smaller than δ, as required. □

We define $T_0=0$ and $T_k=\min\{n>T_{k-1}\colon X_n=i\}$, for $k\ge1$. Recall that $\mathbb{E}_iL^j([T_{k-1},T_k))=1/m_i$ and hence, by the strong Markov property, the random variables $\xi_k:=1-m_i\,L^\nu([T_{k-1},T_k))$ are independent and identically distributed with mean zero. By Lemma 4.1 (a) the first passage time of zero for this walk satisfies $\mathbb{E}_i[N^{1/2}]=\infty$. As $m_iL^i([0,T_*])\ge N-1$ the result follows.

4.2 Proof of Theorem 3 (ii). We first prove the result in the simple case that the state space $\mathcal{S}$ is finite. In this case the chain is positive recurrent and we have $a(n)\sim n$.

Lemma 4.4. *Suppose $\mathcal{S}$ is finite. Then, for any $i\in\mathcal{S}$, we have $\mathbb{E}_i\big[T_*^\beta\big]<\infty$, for all $0\le\beta<\frac12$.*

Proof. Let $T_0=0$ and, for $k\in\mathbb{N}$, define $T_k=\min\{n>T_{k-1}\colon X_n=i\}$. Denote by h_{ij} the probability that the chain started in i hits j before returning to i, and observe that irreducibility implies that $h_{ij}>0$. By the strong Markov property we have $m_jL^j[0,T_1]=YZ$ where Y is a Bernoulli variable with mean h_{ij} and Z is an independent geometric with success parameter h_{ji}. Hence $\mathbb{E}_iL^j[0,T_1)=h_{ij}/m_jh_{ji}$, which also equals $1/m_i$. Recalling that $\mathbb{E}[Z^2]\le 2/h_{ji}^2$ we get $\mathbb{E}_i[L^j[0,T_1)^2]\le 2h_{ij}/m_j^2h_{ji}^2$, and hence $L^\nu[0,T_1)$ has finite variance. Define $\xi_k:=1-m_i\,L^\nu([T_{k-1},T_k))$, and observe that $\xi_1,\xi_2,\dots$ are independent and identically distributed variables with mean zero and finite variance. Let $N:=\min\{n\colon\sum_{k=1}^n\xi_k\le0\}$, and observe that $T_*\le T_N$. Fix $\varepsilon>0$ and note that

$$\mathbb{P}_i\big(T_*>n\big)\le\mathbb{P}_i\big(N>\varepsilon n\big)+\mathbb{P}_i\Big(\sum_{k=1}^{\lceil\varepsilon n\rceil}(T_k-T_{k-1})>n\Big).$$

By Lemma 4.1 (b) the first term on the right-hand side is bounded by a constant multiple of $(\varepsilon n)^{-1/2}$. For the second term we note that the random variables $T_1-T_0,T_2-T_1,\dots$ are independent and identically distributed with finite variance. By Chebyshev's inequality we infer that, for sufficiently small $\varepsilon > 0$, the term is bounded by a multiple of $1/n$. Altogether we get that $\mathbb{P}_i(T_* > n)$ is bounded by a constant multiple of $n^{-1/2}$, from which the result follows immediately. $\square$

We return to the general case. The next result, which is an auxiliary step in the proof of Theorem 3 (ii), may be of independent interest. The short proof given here, which does not make any regularity assumptions on the chain, is due to Vitali Wachtel.

Lemma 4.5. *Fix a state $i \in \mathcal{S}$ and let $T = \min\{n > 0\colon X_n = i\}$ be the first return time to this state. Then*

$$\mathbb{E}_i\big[a_{ii}(T)^\alpha\big] < \infty, \qquad \text{for all } 0 \le \alpha < 1.$$

Proof. By Lemma 1 in Erickson [7], we have for $m(n) := \int_0^n \mathbb{P}_i(T > x)\, dx$ that

$$\frac{n}{m(n)} \le m_i a_{ii}(n) \le 2\,\frac{n}{m(n)} \qquad \text{for all positive integers } n.$$

As $m(n) \ge n\mathbb{P}_i(T > n-1)$ we infer that $m_i a_{ii}(n) \le 2/\mathbb{P}_i(T > n-1)$ and therefore

$$\mathbb{E}_i\big[a_{ii}(T)^\alpha\big] \le \big(\tfrac{2}{m_i}\big)^\alpha \sum_{n=1}^\infty \big(\mathbb{P}_i(T > n-1)\big)^{-\alpha}\, \mathbb{P}_i(T = n) = \big(\tfrac{2}{m_i}\big)^\alpha \sum_{n=1}^\infty (1 - s_{n-1})^{-\alpha}\big(s_n - s_{n-1}\big),$$

where $s_n := \mathbb{P}_i(T \le n)$. Letting $s(t) := s_{n-1} + (t-(n-1))(s_n - s_{n-1})$, for $n-1 \le t < n$, we can bound the sum by $\int_0^\infty (1-s(t))^{-\alpha}\, ds(t)$, which is finite for all $0 \le \alpha < 1$, as required. $\square$

We now look at the reduction of our Markov chain to the finite state space $\mathcal{S}' = \{0\} \cup \{j \in \mathcal{S}\colon \nu_j > 0\}$. More explicitly, let $t_0 = 0$ and $t_k = \min\{n > t_{k-1}\colon X_n \in \mathcal{S}'\}$ for $k \in \mathbb{N}$, and $t_k = \max\{n < t_{k+1}\colon X_n \in \mathcal{S}'\}$ for $k \in -\mathbb{N}$. Then $Y_n = X_{t_n}$ defines an irreducible Markov chain $Y = (Y_n\colon n \in \mathbb{Z})$ with finite state space $\mathcal{S}'$, and its invariant measure is $(m_i\colon i \in \mathcal{S}')$. If N is the stopping time constructed in Theorem 2 for the reduced chain Y, then the solution T_* for the original problem is $T_* = t_N$.

Given two states $i, j \in \mathcal{S}'$ we denote by S_{ij} a random variable whose law is given by $\mathsf{P}(S_{ij} = s) = \mathbb{P}_i(t_1 = s \,|\, Y_1 = j)$ for all $s \in \mathbb{N}$, if $\mathbb{P}_i(Y_1 = j) > 0$, and $S_{ij} = 0$ otherwise. We construct a probability space on which there are independent families $(S_{ij}, S_{ij}^{(k)}\colon k \in \mathbb{N})$ of independent random variables with this law, together with an independent copy of Y and hence N. We denote probability and expectation on this space by P, resp. E. Observe that on this space we can also define a copy of the process $(t_k\colon k \in \mathbb{N})$ by $t_0 = 0$ and

$$t_k = t_{k-1} + \sum_{i,j \in \mathcal{S}'} S_{ij}^{(k)} 1\{Y_{k-1} = i, Y_k = j\} \qquad \text{for } k \in \mathbb{N}.$$

For any non-decreasing, subadditive representative a of the class of the asymptotic Green's function,

$$\mathbb{E}_i\big[a(T_*)^\beta\big] = \mathsf{E}\Big[a\Big(\sum_{k=1}^N t_k - t_{k-1}\Big)^\beta\Big] \le \mathsf{E}\Big[a\Big(\sum_{k=1}^N \sum_{i,j \in \mathcal{S}'} S_{ij}^{(k)}\Big)^\beta\Big] \le \sum_{i,j \in \mathcal{S}'} \mathsf{E}\Big[a\Big(\sum_{k=1}^N S_{ij}^{(k)}\Big)^\beta\Big].$$

It therefore suffices to show that

$$\mathsf{E}\Big[a_{ii}\Big(\sum_{k=1}^N S_{ij}^{(k)}\Big)^\beta\Big] < \infty.$$

Let $n \in \mathbb{N}$ and use first subadditivity of a_{ii} and then Jensen's inequality to get, for $2\beta < \alpha < 1$, that

$$\mathsf{E}\Big[a_{ii}\Big(\sum_{k=1}^n S_{ij}^{(k)}\Big)^\beta\Big] \le \mathsf{E}\Big[\Big(\sum_{k=1}^n a_{ii}^\alpha\big(S_{ij}^{(k)}\big)\Big)^{\beta/\alpha}\Big] \le \Big(\sum_{k=1}^n \mathsf{E}\big[a_{ii}^\alpha\big(S_{ij}^{(k)}\big)\big]\Big)^{\beta/\alpha} = n^{\beta/\alpha}\mathsf{E}\big[a_{ii}^\alpha\big(S_{ij}\big)\big]^{\beta/\alpha}.$$

We now note that, if T_{ij} denotes the first hitting time of state j for X under $\mathbb{P}_i$, we have $\mathsf{P}(S_{ij} > x) \leq C_0\,\mathbb{P}_i(T_{ij} > x)$ for all $x > 0$, where C_0 is the maximum of the inverse of all nonzero transition probabilities from i to all other states, by the chain Y. Hence

$$\mathsf{E}\big[a_{ii}^{\alpha}(S_{ij})\big] \leq C_0\,\mathbb{E}_i\big[a_{ii}^{\alpha}(T_{ij})\big].$$

In the case $i = j$ the right hand side is finite by Lemma 4.5 and, as a_{ii} grows no faster than linearly, the right hand side is finite for all choices of $i, j \in \mathcal{S}'$ by application of Theorem 1.1 in Aurzada et al. [1]. Summarising, we have found a constant $C > 0$ such that

$$\mathsf{E}\Big[a_{ii}\Big(\sum_{k=1}^{n} S_{ij}^{(k)}\Big)^{\beta}\Big] \leq Cn^{\beta/\alpha}.$$

Using the independence of N and $(S_{ij}^{(k)} : k \in \mathbb{N})$ and Lemma 4.4 we get

$$\mathsf{E}\Big[a_{ii}\Big(\sum_{k=1}^{N} S_{ij}^{(k)}\Big)^{\beta}\Big] \leq C\mathbb{E}_i\big[N^{\beta/\alpha}\big] < \infty,$$

as required.

5. Optimality of T_*: Proof of Theorem 4

In this section we prove Theorem 4. We start by introducing an intuitive and convenient way to talk about allocation rules. A path of the Markov chain X can be viewed as leaving white and couloured balls on the integers, in the following way: At each site $k \in \mathbb{Z}$ we place one white ball if $X_k = i$, and $\frac{m_i}{m_j}\nu_j$ balls of colour j if $X_k = j$. By our assumption there is always an integer number of balls at each site. We call a bijection from the set of white balls to the set of coloured balls a *matching*. Given a matching we define an allocation rule $\tau : \Omega \times \mathbb{Z} \to \mathbb{Z}$ by letting

- $\tau(k) = k$ if there is no white ball at site k,
- $\tau(k) = \ell$ if the white ball at site k is matched to a coloured ball at site ℓ.

Every allocation rule thus constructed balances L^{μ} and L^{ν}, for $\mu = \delta_i$. Conversely, every balancing allocation rule agrees L^{μ}-almost everywhere with an allocation rule constructed from a matching. We denote by $\tau_* : \Omega \times \mathbb{Z} \to \mathbb{Z}$ the allocation rule associated with T_* constructed in Proposition 3.1.

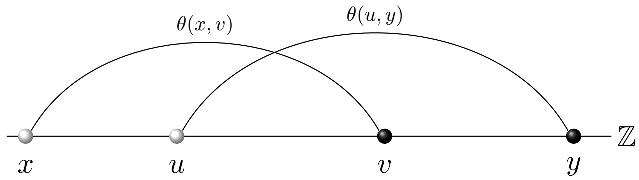
The allocation rule τ_* is associated with the following one-sided *stable matching* or *greedy algorithm*, which is a variant of the famous Gale–Shapley stable marriage algorithm [9].

(1) If the next occupied site to the right of a white ball carries one or more coloured balls, map the white ball to one of those coloured balls.
(2) Remove all white and coloured balls used in step (1) and repeat.

By Lemma 3.3 the algorithm matches every ball after a finite number of steps, and it is easy to see that this leads to the allocation rule τ_*.

Now recall from Section 2 that non-negative, possibly randomized, times T are associated to transport rules $\theta : \Omega \times \mathbb{Z} \times \mathbb{Z} \to [0, 1]$ balancing L^{μ} and L^{ν} with the property that $\theta_\omega(x, y) = 0$ whenever $x > y$. Without loss of generality we may assume that $\theta_\omega(x, x) = 1$ if the site x does not carry a white ball. This implies that, for $x < y$, we can have $\theta_\omega(x, y) > 0$ only if the site x carries a white ball, and the site y carries a coloured ball. Moreover, if y carries a ball of colour j, we have

$$\sum_{x<y} \theta_\omega(x, y) = \frac{m_i}{m_j}\nu_j.$$



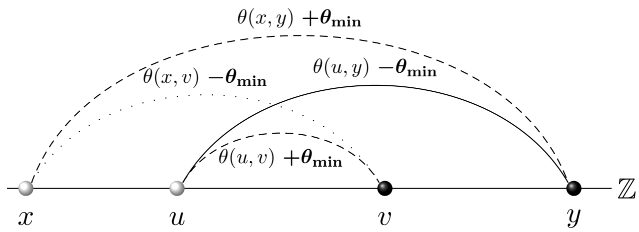


FIGURE 1. The picture above shows a crossing. Its weight $\theta_{\min} := \theta(x,v) \wedge \theta(u,y)$ is assumed to be $\theta(x,v)$, so that in the picture below we see that after the repair the dotted edge has weight zero, and the crossing is therefore removed.

Suppose that $u, v \in \mathbb{Z}$ with $u < v$. We say that the pair (u,v) is *crossed* by θ if there exist sites $x < u < v < y$ such that $\theta(x,v) > 0$ and $\theta(u,y) > 0$. In this case (x,u,v,y) is called a *crossing*.

For a transport rule θ we repair the crossing (x,u,v,y) by letting

- $\theta'(x,y) = \theta(x,y) + (\theta(x,v) \wedge \theta(u,y))$,
- $\theta'(u,v) = \theta(u,v) + (\theta(x,v) \wedge \theta(u,y))$,
- $\theta'(x,v) = \theta(x,v) - (\theta(x,v) \wedge \theta(u,y))$,
- $\theta'(u,y) = \theta(u,y) - (\theta(x,v) \wedge \theta(u,y))$,

and setting $\theta'(w,z) = \theta(w,z)$ if $w \notin \{x,u\}$ or $z \notin \{y,v\}$, see Figure 1. Note that θ' is still a transport rule, the crossing has been repaired, i.e. (x,u,v,y) is not a crossing by θ', and if θ balances L^μ and L^ν then so does θ'.

We now explain how to repair a pair (u,v) crossed by θ by sequentially repairing its crossings and taking limits, so that (u,v) is not crossed by the limiting transport rule. For this purpose we define that a sequence of transport rules θ_n *converges uniformly* to a transport rule θ if

$$\lim_{n\to\infty} \sum_{x,y\in\mathbb{Z}} \big|\theta_n(x,y) - \theta(x,y)\big| = 0.$$

Denote by $y_1, y_2, \ldots$ the sequence of sites $v < y_1 < y_2 < \cdots$ such that $\theta(u,y_n) > 0$, and by $x_1, x_2, \ldots$ the sequence of sites $u > x_1 > x_2 > \cdots$ such that $\theta(x_n,v) > 0$. Note that both sequences could be finite or infinite. First we successively repair the crossings $x_1 < u < v < y_n$, for $n = 1, 2, \ldots$. The total mass moved in the nth repair is bounded by $4\theta(u,y_n)$ and because $\sum_n \theta(u,y_n) \leq 1$ we can infer that the sequence of repaired transport rules converges uniformly to a transport rule θ_1. Of course, here and below if a sequence is finite we take the last element of the sequence as limit. We continue by repairing the crossings $x_2 < u < v < y_n$ of θ_1, for $n = 1, 2, \ldots$, obtaining θ_2, and so on. We obtain a sequence $\theta_1, \theta_2, \ldots$ of transport rules. The amount of mass moved when going from θ_{n-1} to θ_n is bounded by $4\theta(x_n,v)$. As $\sum_n \theta(x_n,v) < \infty$, we infer that the sequence $(\theta_n)_n$ converges uniformly to a limiting transport rule. We observe that this transport rule balances L^μ and L^ν and that (u,v) is not crossed by it.

Lemma 5.1. *Suppose that θ is a transport rule balancing L^μ and L^ν and $A \subset \mathbb{Z}$ a finite interval. Then, by repairing pairs crossed by θ in a given order, we obtain a transport rule θ_* balancing L^μ and L^ν, such that if $u, v \in A$ then (u, v) is not crossed by θ_*.*

Proof. Without loss of generality the left endpoint of A carries a white ball, and its right endpoint carries a coloured ball. Let $v_1, \ldots, v_n$ be the sites in A carrying coloured balls, ordered from left to right. We go through these sites in order, starting with v_1. Take u_1 to be the rightmost site to the left of v_1 carrying a white ball. Repair the pair (u_1, v_1) as above, and observe that the resulting transport rule transports a unit mass from u_1 to v_1. We declare the white ball at site u_1 and one of the coloured balls at v_1 *cancelled*. If v_1 carries an uncancelled ball and there are uncancelled white balls on sites of A to the left of v_1, we choose the rightmost of those, say u_2, repair the pair (u_2, v_1), and cancel two balls as above. We continue until we run out of uncancelled balls. The resulting transport rule has the property that none of the pairs (u, v_1), with $u \in A$, is crossed, and from all sites carrying cancelled white balls a unit mass is transported to site v_1.

We now move to the next coloured ball v_2 and repair all pairs (u, v_2), where u goes from right to left through all sites in $A \cap (-\infty, v_2)$ carrying uncancelled white balls. We do this until we run out of uncancelled white balls to the left of, or coloured balls on the site v_2. Observe that at the end of this step none of the pairs (u, v_1) or (u, v_2), with $u \in A$, is crossed by the resulting transport rule. We continue, moving to the next coloured ball until all coloured balls in A are exhausted. At the end of this finite procedure we obtain a transport rule θ_* balancing L^μ and L^ν, such that if $u, v \in A$ then (u, v) is not crossed by θ_*. □

We call a set A an *excursion* if it is an interval $[m, n]$ such that that there is the same number of white and coloured balls on the sites of A, but the number of white balls exceeds the number of coloured balls on every subinterval $[m, k]$, for $m \leq k < n$. Observe that if A is an excursion, then it is an interval of the form $[m, \tau_*(m)]$ where m carries a white ball, but not all such intervals are excursions. Moreover, for every $x \in A$, we have both $\tau_*(x) \in A$ and $\tau_*^{-1}(x) \subset A$.

Lemma 5.2. *Let A be an excursion and θ_* a transport rule balancing L^μ and L^ν, such that any pair (u, v) with $u, v \in A$ is not crossed by θ_*. Then θ_* agrees in A with the allocation rule τ_*, in the sense that $\theta_*(x, y) = 1\{\tau_*(x) = y\}$ and $\theta_*(y, x) = 1\{\tau_*(y) = x\}$, for all $x \in A$ and $y \in \mathbb{Z}$.*

Proof. We start by fixing a site $x \in A$ carrying a white ball, and note that, by definition of an excursion, we also have $\tau_*(x) \in A$. We show by contradiction that θ_* transports no mass from x to a point other than $\tau^*(x)$.

First, suppose that there exist $x < v < \tau_*(x)$ with $\theta_*(x, v) > 0$. As there are more white than coloured balls on the sites in $[x, v]$, and as every site carries at most one white ball, we find $x' \in (x, v)$ such that the sites of $[x', v]$ carry the same number of white and coloured balls. As $\theta_*(x, v) > 0$ not all white balls in $[x', v]$ are matched within that interval, and there must also exist $u \in [x', v)$ and $y > v$ such that $\theta_*(u, y) > 0$. So we have found a pair (u, v) with $u, v \in A$, which is crossed by θ_*, and hence a contradiction.

Second, suppose that there exist $v > \tau_*(x)$ with $\theta_*(x, v) > 0$. As there are at least as many coloured balls as white balls in $[x, \tau_*(x)]$ not all coloured balls are matched within that interval, and hence there exists a $y \in (x, \tau_*(x)]$ and a site $u < x$ with $\theta_*(u, y) > 0$. So we have found a pair (x, y) with $x, y \in A$, which is crossed by θ_*, and hence a contradiction. We conclude that $\theta_*(x, y) = 1\{\tau_*(x) = y\}$ for all $x \in A$.

Now fix a site $x \in A$ carrying balls of colour j. Then $\tau_*^{-1}(x)$ is a set of $(m_i/m_j)\nu_j$ points in A. Hence, by the first part, $\theta_*(y,x) = 1\{\tau_*(y) = x\}$ for all $y \in \tau_*^{-1}(x)$. Moreover,

$$\sum_{y \in \tau_*^{-1}(x)} \theta_*(y,x) = (m_i/m_j)\nu_j = \sum_{y \in \mathbb{Z}} \theta_*(y,x).$$

Hence $\theta_*(y,x) = 0 = 1\{\tau_*(y) = x\}$ also for all $y \notin \tau_*^{-1}(x)$. $\square$

We now let ψ be a non-negative, concave function on the non-negative integers $\mathbb{N}_0$. Note that this implies that $\psi\colon \mathbb{N}_0 \to [0,\infty)$ is non-decreasing. We further assume that $\psi(0) = 0$, an assumption which causes no loss of generality in Theorem 4. We write $\psi(n) = 0$ for $n \le 0$ to simplify the notation.

Lemma 5.3. *Let A be an excursion and suppose θ is a transport rule balancing L^μ and L^ν. Then*

$$\sum_{x \in A} \psi\big(\tau_*(x) - x\big) + \sum_{x \in \tau_*^{-1}(A)} \psi\big(\tau_*(x) - x\big) \le \sum_{\substack{x \in A\\ y \in \mathbb{Z}}} \theta(x,y)\, \psi(y-x) + \sum_{\substack{x \in \mathbb{Z}\\ y \in A}} \theta(x,y)\, \psi(y-x).$$

Proof. Observe that, by concavity, for all $a, b, c \in \mathbb{N}_0$, we have

$$\psi(a+b) + \psi(b+c) \ge \psi(a+b+c) + \psi(b). \tag{5.1}$$

Fix a crossing $x < u < v < y$ with $u, v \in A$, and let θ' be the result of repairing the crossing. We show that repairing the crossing does not increase

$$\sum_{\substack{x \in A\\ y \in \mathbb{Z}}} \theta(x,y)\, \psi(y-x) + \sum_{\substack{x \in \mathbb{Z}\\ y \in A}} \theta(x,y)\, \psi(y-x)$$

by looking at the difference of the repaired and original state of the sum. If $x, y \notin A$ we get

$$\begin{aligned}
&\theta'(u,y)\psi(y-u) + 2\theta'(u,v)\psi(v-u) + \theta'(x,v)\psi(v-x)\\
&\qquad - \big(\theta(u,y)\psi(y-u) + 2\theta(u,v)\psi(v-u) + \theta(x,v)\psi(v-x)\big)\\
&\quad = \big(\theta(x,v) \wedge \theta(u,y)\big)\big(2\psi(v-u) - \psi(v-x) - \psi(y-u)\big) \le 0,
\end{aligned}$$

as ψ is non-decreasing. If $x \in A, y \notin A$ we get

$$\begin{aligned}
&\theta'(u,y)\psi(y-u) + 2\theta'(u,v)\psi(v-u) + 2\theta'(x,v)\psi(v-x) + \theta'(x,y)\psi(y-x)\\
&\qquad - \big(\theta(u,y)\psi(y-u) + 2\theta(u,v)\psi(v-u) + 2\theta(x,v)\psi(v-x) + \theta(x,y)\psi(y-x)\big)\\
&\quad = \big(\theta(x,v) \wedge \theta(u,y)\big)\big(2\psi(v-u) + \psi(y-x) - 2\psi(v-x) - \psi(y-u)\big)\\
&\quad \le \big(\theta(x,v) \wedge \theta(u,y)\big)\big(\psi(v-u) + \psi(y-x) - \psi(v-x) - \psi(y-u)\big) \le 0,
\end{aligned}$$

using first that ψ is non-decreasing and then (5.1). The case $x \notin A, y \in A$ is analogous. If $x, y \in A$ the difference is twice

$$\begin{aligned}
&\theta'(x,v)\psi(v-x) + \theta'(u,y)\psi(y-u) + \theta'(u,v)\psi(v-u) + \theta'(x,y)\psi(y-x)\\
&\qquad - \big(\theta(x,v)\psi(v-x) + \theta(u,y)\psi(y-u) + \theta(u,v)\psi(v-u) + \theta(x,y)\psi(y-x)\big)\\
&\quad = \big(\theta(x,v) \wedge \theta(u,y)\big)\big(\psi(y-x) + \psi(v-u) - \psi(v-x) - \psi(y-u)\big) \le 0,
\end{aligned}$$

by application of (5.1), which shows that in all cases the sum above is not increased by the repair.

Repairing crossings successively as described in Lemma 5.1, we get

$$\sum_{\substack{x \in A\\ y \in \mathbb{Z}}} \theta_*(x,y)\, \psi(y-x) + \sum_{\substack{x \in \mathbb{Z}\\ y \in A}} \theta_*(x,y)\, \psi(y-x) \le \sum_{\substack{x \in A\\ y \in \mathbb{Z}}} \theta(x,y)\, \psi(y-x) + \sum_{\substack{x \in \mathbb{Z}\\ y \in A}} \theta(x,y)\, \psi(y-x).$$

By Lemma 5.2 we have $\theta_*(x,y)=1\{\tau(x)=y\}$ if $x\in A$ or $y\in A$, and this allows us to rewrite the left hand side as stated. $\square$

Lemma 5.4. *Let $T\geq 0$ be a (possibly randomized) unbiased shift and $\theta\colon \Omega\times\mathbb{Z}\times\mathbb{Z}\to[0,1]$ be the associated transport rule. Denote by $(T_n\colon n\in\mathbb{Z})$ the times in which X visits the state i, in order so that $T_0=0$. Let $\psi\colon\mathbb{Z}\to[0,\infty)$ be concave. Then, $\mathbb{P}_i$-almost surely,*

$$\lim_{n\to\infty}\frac{1}{n}\Big\{\sum_{k=0}^{T_n-1}\sum_{\ell=k+1}^{\infty}\theta(k,\ell)\psi(\ell-k)+\sum_{\ell=-\infty}^{k-1}\theta(\ell,k)\psi(k-\ell)\Big\}=2\mathbb{E}_i^{\oplus}\psi(T),$$

and

$$\lim_{m\to\infty}\frac{1}{m}\Big\{\sum_{k=T_{-m}}^{1}\sum_{\ell=k+1}^{\infty}\theta(k,\ell)\psi(\ell-k)+\sum_{\ell=-\infty}^{k-1}\theta(\ell,k)\psi(k-\ell)\Big\}=2\mathbb{E}_i^{\oplus}\psi(T).$$

Proof. We observe, from the strong Markov property, that $\xi_n=(X_{T_{n-1}+1},\dots,X_{T_n})$, $n\in\mathbb{Z}$, are independent and identically distributed random vectors. Hence their shift is stationary and ergodic, see for example [6, 8.4.5]. By the ergodic theorem, see e.g. [6, 8.4.1], $\mathbb{P}_i$-almost surely,

$$\lim_{n\to\infty}\frac{1}{n}\Big\{\sum_{k=0}^{T_n-1}\sum_{\ell=k+1}^{\infty}\theta(k,\ell)\psi(\ell-k)\Big\}=\mathbb{E}_i\sum_{\ell=1}^{\infty}\theta_\omega(0,\ell)\psi(\ell)=\mathbb{E}_i^{\oplus}\psi(T).$$

Similarly,

$$\lim_{n\to\infty}\frac{1}{n}\Big\{\sum_{k=0}^{T_n-1}\sum_{\ell=-\infty}^{k-1}\theta(\ell,k)\psi(k-\ell)\Big\}=\mathbb{E}_i\sum_{k=0}^{T_1-1}\sum_{\ell=-\infty}^{k-1}\theta_\omega(\ell,k)\psi(k-\ell).$$

The expectation equals

$$\sum_{j\in\mathcal{S}}\frac{m_j}{m_i}\mathbb{E}_j\sum_{\ell=-\infty}^{-1}\theta_\omega(\ell,0)\psi(-\ell)=\frac{1}{m_i}\mathbb{E}\sum_{\ell=-\infty}^{-1}\theta_\omega(\ell,0)\psi(-\ell)=\frac{1}{m_i}\mathbb{E}\sum_{\ell=1}^{\infty}\theta_\omega(0,\ell)\psi(\ell)=\mathbb{E}_i^{\oplus}\psi(T),$$

using translation invariance of θ. The second statement follows in the same manner. $\square$

Proof of Theorem 4. We now look at the sequence

$$\tau_n=\min\big\{T_k\geq 0\colon L^\mu([0,T_k))-L^\nu([0,T_k))\leq \tfrac{-n}{m_i}\big\}.$$

Let $d_n=L^\mu([0,\tau_n))-L^\nu([0,\tau_n))$ and define

$$\sigma_n=\max\big\{k\leq 0\colon -L^\mu([k,0))+L^\nu([k,0))=d_n\big\}.$$

(τ_n) and (σ_n) are well-defined subsequences of $(T_n\colon n\in\mathbb{Z})$, $\mathbb{P}_i$-almost surely, by Lemma 3.3. Moreover, $\tau_n\uparrow\infty$, $\sigma_n\downarrow-\infty$ and by construction $[\sigma_n,\tau_n-1]$ is an excursion, see Figure 2. By Lemma 5.3

$$\sum_{k=\sigma_n}^{\tau_n-1}\Big\{\psi\big(\tau_*(k)-k\big)+\sum_{\ell\in\tau_*^{-1}(k)}\psi\big(\tau_*(\ell)-\ell\big)\Big\}\leq\sum_{\substack{\sigma_n\leq k\leq\tau_n-1\\ \ell\in\mathbb{Z}}}\theta(k,\ell)\psi(k-\ell)+\sum_{\substack{\sigma_n\leq\ell\leq\tau_n-1\\ k\in\mathbb{Z}}}\theta(k,\ell)\psi(k-\ell).$$

Lemma 5.4 shows that the left hand side is asymptotically equivalent to $2m_i\,L^i([\sigma_n,\tau_n])\,\mathbb{E}_i\psi(T_*)$ and the right hand side to $2m_i\,L^i([\sigma_n,\tau_n])\,\mathbb{E}_i^{\oplus}\psi(T)$, from which we conclude that $\mathbb{E}_i\psi(T_*)\leq\mathbb{E}_i^{\oplus}\psi(T)$. $\square$

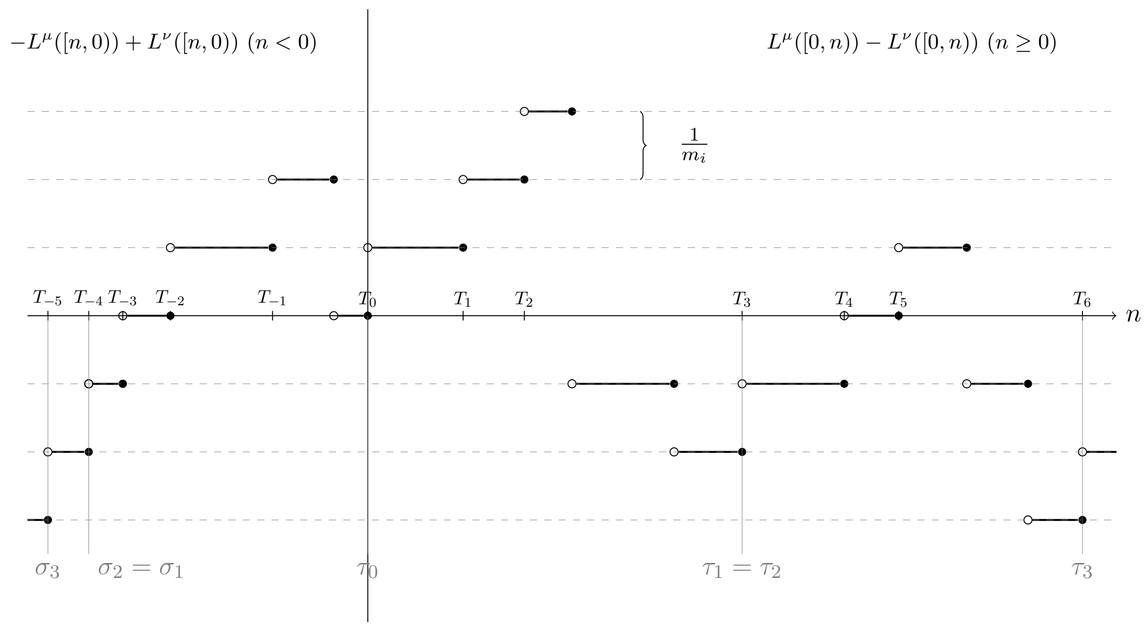


FIGURE 2. A possible profile of local time differences over the excursion $[\sigma_3, \tau_3 - 1]$. Upward jumps are of size $1/m_i$, downward jumps are a positive integer multiple of $1/m_i$, the actual value depending on the colour of the ball at the location of the jump.

6. Concluding remarks and open problems

Non-Markovian setting. Theorem 1 and Theorem 2 remain valid in a more general non-Markovian setting. We require that under the σ-finite measure $\mathbb{P}$ the stochastic process X, taking values in the countable state space $\mathcal{S}$, is stationary with a strictly positive stationary σ-finite measure $(m_i : i \in \mathcal{S})$. The probability measure $\mathbb{P}_i$ is then defined by conditioning X on the event $\{X_0 = i\}$. We further require that, for every $i, j \in \mathcal{S}$, the random sets $\{n \in \mathbb{N} \colon X_n = j\}$ and $\{n \in \mathbb{N} \colon X_{-n} = j\}$ are infinite $\mathbb{P}_i$-almost surely. Then both theorems carry over to this conditioned process. Further technical conditions are required to generalize Lemma 5.4 and hence extend Theorem 4 to the non-Markovian setting. Theorem 3 however fully exploits the Markov structure and cannot be generalized easily.

General inital distribution. Although our main focus is on the case where the initial distribution is the Dirac measure δ_i for some $i \in \mathcal{S}$, the statements of Proposition 2.1 and 2.2 allow general initial distributions μ. By conditioning on the initial state one can see that a *sufficient* condition for existence of the solution is that the target measure ν admits a decomposition $\nu = \sum_{i \in \mathcal{S}} \nu^{(i)} \mu_i$, where $\nu^{(i)}$ are probability measures on $\mathcal{S}$, such that $m_i \nu_j^{(i)}/m_j$ are integers for all $i, j \in \mathcal{S}$. We do not believe that this is also a *necessary* condition.

Randomized shifts. If the target measure ν fails to satisfy the integer condition in Theorem 1 (b), extra randomization is needed to solve the embedding problem. With extra randomness any target measure ν may be embedded in a way similar to the extra head schemes in [12]: Take a random variable $U \sim \text{Uniform}(0,1)$ and define

$$T_{\texttt{rand}} := \min\Big\{ n \geq 0 \colon L^i([0,n]) - \sum_{j \in \mathcal{S}} \nu_j \, L^j([0,n]) \leq \tfrac{U}{m_i} \Big\}. \tag{6.1}$$

Then $T_{\texttt{rand}}$ is an unbiased shift embedding ν. We see that if the integer condition holds, the sample value of U becomes irrelevant and we recover the non-randomized solution T_* defined in Theorem 2.

Brownian motion and optimal shifts. Last et al. [13] discuss the Skorokhod embedding problem for a two-sided Brownian motion $(B_t)_{t\in\mathbb{R}}$. In this context a random time T solves the embedding problem if $(B_{T+t}-B_T)_{t\in\mathbb{R}}$ is a standard two-sided Brownian motion independent of B_T and the law of B_T is ν. They show that for *any* target distribution ν not charging the origin the stopping time $T_* = \inf\{t>0\colon L_t^0 = L_t^\nu\}$, where $(L_t^x\colon t>0)$ is the process of local times at level x and $L_t^\nu := \int L_t^x\,\nu(dx)$, solves the embedding problem. They further show that every solution T that is a stopping time satisfies $\mathbb{E}[T^{\frac14}]=\infty$ while under a mild condition on ν the constructed solution T_* satisfies $\mathbb{E}[T_*^\beta]<\infty$ for all $\beta<\frac14$. The techniques of the present paper can be adapted to improve the results of [13] by showing that $\mathbb{E}[T^{\frac14}]=\infty$ even for non-negative solutions which are not necessarily stopping times, and also to show a strong optimality result similar to Theorem 4, i.e. that $\mathbb{E}_0\psi(T_*)\le\mathbb{E}_0\psi(T)$ simultaneously for all non-negative concave functions ψ. These results will appear in the forthcoming thesis [17].

Signed shifts. The optimality result of Theorem 4 cannot be extended easily to random times T that can take both positive and negative values. Indeed, starting from such a solution T and associating an allocation rule τ to it, we may still make local improvements by repairing crossings, but now there is more than one way to repair a crossing and the optimal way to do this appears to involve nonlocal choices. To get a feeling for the difficulties, we look at a two-sided stable matching strategy that at a first glance looks like a good candidate for an optimal solution. In the language of Section 5 we match a coloured ball to a white ball if *both* the coloured ball is the nearest coloured ball to the white ball, and the white ball is the nearest white ball to the coloured ball (resolving possible ties in some deterministic way). Locally, the resulting allocation rule may be better or worse than the one coming from our one-sided stable matching. Consider, for example, configuration of balls in the order white–coloured–white–coloured placed at distances a,b,c such that $b<a,c$. The two-sided algorithm matches the middle balls and, if other balls are sufficiently far away, the outer balls, which gives a contribution of $\psi(b)+\psi(a+b+c)$. One-sided stable matching matches the first pair and the second pair and gives $\psi(a)+\psi(c)$, and each contribution could be smaller or larger depending on the relative size of a,b,c. Even finding the optimal moment properties of signed shifts is an open problem.

Random fields. A vast open area of possible further research are embedding problems for multi-parameter processes and random fields. In higher dimensions stable allocation procedures no longer have optimal moment properties, see for example Holroyd, Peres and Schramm [11], so other methods need to be considered. It would be particularly interesting to investigate embedding problems for spin systems such as the infinite volume Gibbs measure of the Ising model at high temperature.

Acknowledgments: We would like to thank Vitali Wachtel for valuable discussions on Wiener-Hopf decompositions and for suggesting the proof of Lemma 4.5, which allowed us to remove a regularity condition on the asymptotic Green's function of the Markov chain. The first author would like to thank Mathias Beiglböck and Martin Huesmann for enlightening discussions during visits to the Hausdorff Institute and Eurandom. We would particularly like to thank Martin Huesmann for conjecturing the result of Theorem 4. Last but not least, we would like to thank Günter Last and an anonymous referee for several insightful comments.

References

[1] F. Aurzada, H. Döring, M. Ortgiese, and M. Scheutzow. Moments of recurrence times for Markov chains. *Electron. Commun. Probab.*, 16:296–303, 2011.

[2] M. T. Barlow and M. Yor. (Semi-) martingale inequalities and local times. *Z. Wahrsch. Verw. Gebiete*, 55:237–254, 1981.

[3] R. F. Bass. *Probabilistic techniques in analysis.* Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[4] M. Beiglböck, A.M.G. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *arXiv:1307.3656*, 2013.

[5] X. Chen. How often does a Harris recurrent Markov chain recur? *Ann. Probab.*, 27:1324–1346, 1999.

[6] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002.

[7] K. B. Erickson. The strong law of large numbers when the mean is undefined. *Trans. Amer. Math. Soc.*, 185:371–381, 1973.

[8] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[9] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *Amer. Math. Monthly*, 69:9–15, 1962.

[10] 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.

[11] A. E. Holroyd, R. Pemantle, Y. Peres, and O. Schramm. Poisson matching. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45:266–287, 2009.

[12] A. E. Holroyd and Y. Peres. Extra heads and invariant allocations. *Ann. Probab.*, 33:31–52, 2005.

[13] G. Last, P. Mörters, and H. Thorisson. Unbiased shifts of Brownian motion. *Ann. Probab.*, 42:431–463, 2014.

[14] G. Last and H. Thorisson. Invariant transports of stationary random measures and mass-stationarity. *Ann. Probab.*, 37:790–813, 2009.

[15] 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.

[16] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. Oxford University Press, New York, 1995.

[17] I. Redl. *Unbiased shifts of stochastic processes.* PhD thesis, University of Bath, forthcoming.

[18] F. Spitzer. A Tauberian theorem and its probability interpretation. *Trans. Amer. Math. Soc.*, 94:150–169, 1960.

[19] H. Thorisson. *Coupling, Stationarity, and Regeneration.* Probability and its Applications (New York). Springer-Verlag, New York, 2000.