CENTRAL LIMIT THEOREM FOR EXCITED RANDOM WALK IN THE RECURRENT REGIME

DMITRY DOLGOPYAT

Abstract. We consider excited random walk on a strip. We assume that the cookies are positive and that the total expected drift per site is less than \(1/L\) where \(L\) is the width of the strip. We prove a quenched limit theorem claiming that the position of the walker converges after the diffusive rescaling to a perturbed Brownian Motion.

Let \(Y = \mathbb{Z} \times (\mathbb{Z}/L\mathbb{Z})\), where \(L > 1\) is an integer, \(G = \{-e_1, e_1, -e_2, e_2\}\) where \(e_j\) are coordinate vectors. We denote the coordinates of points \(y \in Y\) by \((x(y), s(y))\). Consider a cookie environment on \(Y\), that is, for each \(y \in Y\), \(j \in \mathbb{N}\), there is a probability distribution \(\omega(y, j, e)\) on \(G\). Consider an excited random walk \(Y_n = (X_n, S_n)\) that is

\[
P(Y_{n+1} - Y_n = e|Y_1, \ldots, Y_n) = \omega(Y_n, l_n, e)
\]

where \(l_n\) is the number of visits to \(Y_n\) by time \(n\). (We denote by \(P\) and \(E\) the quenched probability and expectation with fixed \(\omega\) and by \(P^0\) and \(E^0\) the annealed probability and expectation.) \(Y_n\) is called (multi-) excited random walk (ERW). We make the following assumptions:

(A) \(\delta(y, j) := \omega(y, j, e_1) - \omega(y, j, -e_1) \geq 0\),
(B) \(\exists \kappa > 0\) such that \(\omega(y, j, e) \geq \kappa\),
(C) \(\omega\) is stationary with respect to \(G\)-shifts and ergodic.
(D) Let \(\delta(y) = \sum_{j=1}^{\infty} \delta(y, j)\) then

\[
\delta := E(\delta(y)) < \frac{1}{L}.
\]

(E) For each \(\varepsilon > 0\) there exists \(N(\varepsilon, y)\) such that for each \(j \geq N\), for each \(e \in G\) \(|\omega(y, j, e) - \frac{1}{4}| < \varepsilon\). Moreover \(E(N(\varepsilon, y)) < \infty\).

The quantity \(\delta\) introduced in (D) plays a crucial role in description of the behavior of ERW. In particular \(Y_n\) is recurrent in the sense that every site is visited infinitely often iff \(\delta L \leq 1\), see [10, 11, 1]. (In case \(\delta L < 1\) which is a subject of the our work recurrence also follows from Lemma 5 of the present paper.) Several papers addressed the limiting

\[1\]
behavior of the ERW in the transient regime \([9, 2, 3, 8, 7]\). Our paper deals with recurrent ERW.

Let \(\mathcal{B}(t)\) denote the Brownian motion with variance \(\frac{1}{2}\). Recall ([5]) that for all \(\alpha, \beta < 1\) and for almost every realization of \(\mathcal{B}\) there exists a unique solution \(\mathcal{W}(t)\) of the equation

\[
\mathcal{W}(t) = \mathcal{B}(t) + \alpha \max_{[0,t]} \mathcal{W}(s) + \beta \min_{[0,t]} \mathcal{W}(s)
\]

which is called \((\alpha, \beta)\)-perturbed Brownian Motion.

Define \(\mathcal{W}_n(t)\) by setting \(\mathcal{W}_n(m/n) = X_m \sqrt{n}\) and interpolating linearly in between.

**Theorem 1.** For almost every \(\omega\), \(\mathcal{W}_n(t)\) converges weakly as \(n \to \infty\) to \((\alpha, \beta)\)-perturbed Brownian Motion where \(\alpha = -\beta = \delta L\).

**Remark.** A similar result is valid for ERW on \(\mathbb{Z}\) with obvious modifications. Namely, \(G = \{-e, +e\}\), condition (E) becomes \(|\omega(y,j,e) - \frac{1}{2}| < \varepsilon\) and the variance of the limiting Brownian Motion equals \(t\).

**Remark.** Our result leaves open the critical case \(\delta L = 1\). (Observe that (1) is not well posed if \(\alpha = 1\).)

We divide the proof into several steps. Fix \(T > 0\).

**Lemma 1.** For any \(m\) there is a constant \(\gamma_m^-\) such that for any \(\omega\) for any stopping time \(\sigma\) for any \(N \in \mathbb{N}\) we have

\[
P\left(\min_{k \leq N} (X_{\sigma+k} - X_{\sigma}) \leq -R\sqrt{N}\right) \leq \frac{\gamma_m^-}{R^{2m}}.
\]

In particular

\[
P\left(\min_{[0,T]} \mathcal{W}_n(t) < -R\right) \leq \frac{\tilde{\gamma}_m^-}{R^{2m}}
\]

where \(\tilde{\gamma}_m = T^m \gamma_m^-\).

**Proof.** Denote

\[
\Delta_k = X_{k+1} - X_k, \quad \bar{\Delta}_k = \mathbb{E}(\Delta_k|Y_1, \ldots, Y_k) = \delta(Y_k, l_k),
\]

\[
C_n = \sum_{k=0}^{n-1} \bar{\Delta}_k, \quad B_n = \sum_{k=0}^{n-1} [\Delta_k - \bar{\Delta}_k].
\]

By assumption (A), \(X_{\sigma+k} - X_{\sigma} \geq B_{\sigma+k} - B_{\sigma}\). Since \(M_k = B_{\sigma+k} - B_{\sigma}\) is a martingale with respect to the \(\sigma\)-algebra generated by \(\Delta_0, \ldots, \Delta_{\sigma+k-1}\) and the quadratic variation of \(M\) grows at most linearly, it follows from [6], Theorem 2.11 that that for each \(m \in \mathbb{N}\) there is a constant \(\gamma_m^-\) such that

\[
\mathbb{E}(\max_{k \leq n} |B_k|^m) \leq \gamma_m^- n^m
\]
and so by Markov inequality

\begin{equation}
\mathbb{P}(\max_{k \leq n} |B_k| \geq R\sqrt{n}) \leq \frac{\gamma_m}{R^{2m}}.
\end{equation}

which implies the result we need. \qed

Denote

\[A_{n_0} = \left\{ \omega : \sum_{x(y) = - \frac{(1 - \delta L)n}{3}}^n \delta(y) < (2 + \delta L)n \text{ for all } n \geq n_0 \right\} .\]

Note that by the Ergodic Theorem

\begin{equation}
\mathbf{P}(A_{n_0}) \rightarrow 1 \text{ as } n_0 \rightarrow \infty.
\end{equation}

Let \(T\) denote the space shift \((T^k \omega)((x, s), j, e) = \omega((x + k, s), j, e)\)

\textbf{Lemma 2.} There is a constant \(\gamma_m^+\) such that for any \(\omega\) such that \(T^\sigma \omega \in A_{n_0}\) for any stopping time \(\sigma\) such that \(X_\sigma = x\) for any numbers \(R \in \mathbb{R}, N \in \mathbb{N}\) such that \(R\sqrt{N} \geq n_0\) we have

\[\mathbb{P}\left(\max_{k \leq N} (X_{\sigma+k} - X_\sigma) \geq R\sqrt{N}\right) \leq \frac{\gamma_m^+}{R^{2m}}.\]

In particular for almost every \(\omega\) we have

\[\mathbb{P}(\max_{[0, T]} \mathcal{W}_n(t) > R) \leq \frac{\gamma_m^+}{R^{2m}}\]

provided that \(n\) is large enough, where \(\gamma_m^+ = T^m \gamma_m^+\).

\textbf{Proof.} Denote

\[\tilde{X}_k = X_{\min(\sigma+k, \hat{\sigma})} - X_\sigma, \quad \tilde{M}_k = M_{\min(k, \sigma - \hat{\sigma})}\]

where \(M\) is the martingale from the proof of Lemma 1 and \(\hat{\sigma}\) is the first time after \(\sigma\) when \(X_\hat{\sigma} = X_\sigma - \left[R\sqrt{N\frac{1 - \delta L}{3}}\right]\). In view of Lemma 1 it suffices to show that

\[\mathbb{P}\left(\max \tilde{X}_k \geq R\sqrt{N}\right) \leq \frac{\gamma_m^+}{R^{2m}}.\]

By the definition of \(A_{n_0}\) we have \(\tilde{X}_k \geq \tilde{M}_k + R\sqrt{N\frac{2 + \delta L}{3}}\) so if \(\tilde{X}_k \geq R\sqrt{N}\) then \(\tilde{M} \geq R\sqrt{N\frac{1 - \delta L}{3}}\). Now the statement of the Lemma follows from (2). \qed

Let \(r_n = \max_{k \leq n}(X_k) - \min_{k \leq n}(X_k)\) denote the range of the walk.

\textbf{Lemma 3.} For almost every \(\omega\) the following holds. As \(n \rightarrow \infty\) the process \(B_n(t) := \frac{B_{|n|}}{\sqrt{n}}\) converges weakly to \(\mathbb{B}(t)\).
Proof. Since $B_n$ is a martingale it suffices, due to [6], Theorem 4.4, to show that for almost every $\omega$

$$\sup_{t \in [0, T]} \left| \frac{V[nt]}{n} - \frac{t}{2} \right| \to 0 \text{ in probability as } n \to \infty$$

where $V_n$ is the quadratic variation of $B_n$. For the discrete time process it is enough to show that for almost every $\omega$

$$\max_{0 \leq m \leq n} \left| \frac{V_m}{n} - \frac{m}{2n} \right| \to 0 \text{ in probability as } n \to \infty.$$ 

Fix $\varepsilon > 0$. Choose $N_0$ such that

$$\mathbf{E}(\mathcal{N}(\varepsilon, y) - N_0) < \varepsilon$$

where $\mathcal{N}(\varepsilon, y)$ is a constant from condition (E). Split $V_m = V_m^- + V_m^+$ where

$$V_m^- = \sum_{k=0}^{m-1} [\Delta_k - \bar{\Delta}_k]^2 I(l_k \leq N_0), \quad V_m^+ = \sum_{k=0}^{m-1} [\Delta_k - \bar{\Delta}_k]^2 I(l_k > N_0).$$

Then $V_m^- \leq 4N_0 Lr_m \ll n$ (by Lemmas 1 and 2) whereas

$$V_m^+ = \frac{m}{2} + \varepsilon'_m + \varepsilon''_m$$

where

$$\varepsilon'_m = \sum_k \left( [\Delta_k - \bar{\Delta}_k]^2 - \frac{1}{2} \right) I(l_k > \max(N(\varepsilon, Y_k), N_0)),$$

$$\varepsilon''_m = \sum_k \left( [\Delta_k - \bar{\Delta}_k]^2 - \frac{1}{2} \right) I(N_0 < l_k \leq N(\varepsilon, Y_k)).$$

Observe that on $l_k > N(\varepsilon, Y_k)$ we have

$$\left| [\Delta_k - \bar{\Delta}_k]^2 - \frac{1}{2} \right| = \left| \omega(Y_k, l_k, e_1) + \omega(Y_k, l_k, -e_1) - \frac{1}{2} \right| - \left[ \omega(Y_k, l_k, e_1) - \omega(Y_k, l_k, -e_1) \right]^2 \leq 2\varepsilon + (2\varepsilon)^2$$

and so $|\varepsilon''_m| \leq (2\varepsilon + (2\varepsilon)^2) n$. On the other hand

$$\left| \mathcal{N}(\varepsilon, y) - N_0 \right|$$

where the summation in (*) runs over $y$ with

$$\min_{k \leq n} (X_k) \leq x(y) \leq \max_{k \leq n} (X_k).$$

So (4) and the ergodic theorem ensure that $|\varepsilon''_m|$ is less than $2\varepsilon Lr_n$ provided that $r_n$ is large enough (if $r_n$ is small then our claim that $|\varepsilon''_m| \ll n$ is obvious). This concludes the proof of Lemma 3. \qed
Lemma 4. \{W_n(t)\} is tight.

Proof. Since \(X_0 = 0\) [4], Lemma 8.3 implies that in order to prove tightness it suffices to show that for almost all \(\omega\) given positive constants \(\varepsilon, \eta\) there exists a positive constant \(\delta\) such that if \(n\) is sufficiently large then for all \(t \leq T\)

\[
\frac{1}{\delta} \mathbb{P}\left( \sup_{s \in [t, t+\delta]} |W_n(s) - W_n(t)| \geq \varepsilon \right) \leq \eta.
\]

Without rescaling this amounts to showing that for all \(n_1 \leq nT\) we have

\[
\frac{1}{\delta} \mathbb{P}\left( \max_{n_1 \leq n_2 \leq n_1 + \delta n} |X_{n_2} - X_{n_1}| \geq \varepsilon \sqrt{n} \right) \leq \eta.
\]

Take \(\delta\) such that

\[
\gamma^{-\delta} < \frac{\varepsilon^{25}}{23} \quad \text{and} \quad \frac{\gamma^{+\delta}}{23} < \eta
\]

By Lemmas 1 and 2 given \(\eta, \delta\) there exists \(R\) such that

\[
\mathbb{P}\left( \max_{k \leq T_n} |X_k| \geq R \sqrt{n} \right) \leq \frac{3}{\delta \eta}
\]

so it suffices to show that

\[
\frac{1}{\delta} \mathbb{P}\left( \max_{n_1 \leq n_2 \leq n_1 + \delta n} |X_{n_2} - X_{n_1}| \geq \varepsilon \sqrt{n} \text{ and } |X_{n_1}| \leq R \sqrt{n} \right) \leq \frac{2\eta}{3}.
\]

We shall show that

\[
\frac{1}{\delta} \mathbb{P}\left( \max_{n_1 \leq n_2 \leq n_1 + \delta n} X_{n_2} \geq X_{n_1} + \varepsilon \sqrt{n} \text{ and } |X_{n_1}| \leq R \sqrt{n} \right) \leq \frac{\eta}{3},
\]

the lower bound on \(X_{n_2}\) is similar. Take \(n_0\) such that \(\mathbb{P}(A_{n_0}^c) \leq \frac{\varepsilon}{100R}\).

Then by the Ergodic Theorem for large \(n\)

\[
\sum_{x = -2R \sqrt{n}}^{2R \sqrt{n}} I_{A_{n_0}^c}(T^x \omega) \leq \frac{2\varepsilon}{25} \sqrt{n}
\]

where \(I\) denotes the indicator function. Hence there exists \(x\) such that \(X_{n_1} \leq x \leq X_{n_1} + \frac{2\varepsilon}{25} \sqrt{n}\) such that \(T^x \omega \in A_{n_0}\). Let \(\sigma\) be the first time after \(n_1\) when \(X_\sigma = x\). Applying Lemma 2 with \(m = 2\) we get

\[
\frac{1}{\delta} \mathbb{P}\left( X_{\sigma + k} - X_\sigma > \frac{23\varepsilon}{25} \sqrt{n} \right) \leq \frac{\gamma^{+\delta}}{23} < \eta
\]

where the last inequality follows from (6). This proves (7) and completes the proof of Lemma 4. \qed
Let 
\[ Z(a, b) = \sum_{(x,s): a \leq x \leq b} \delta(x, s) \]
denote the total amount of cookies stored between \(a\) and \(b\). We shall denote by \(\tau_x\) the first time \(X_\tau = x\). Let 
\[ \tau(x, M) = \begin{cases} 
\tau_{x+M} & \text{if } x \geq 0 \\
\tau_{x-M} & \text{if } x < 0 
\end{cases} \]

The next lemma is a quantitative version of the recurrence results of [10, 11].

**Lemma 5.** For each \(N, \varepsilon\) there exists a number \(M\) and a set \(\Omega_M\) such that \(\mathbb{P}(\Omega_M) > 1 - \varepsilon\) and for all \(x \in \mathbb{Z}\) for all \(\omega\) such that \(T^x\omega \in \Omega_M\) we have for each \(s \in \mathbb{Z}/L\mathbb{Z}\)

\[ (8) \quad \mathbb{P}(Y_n \text{ visits } (x, s) \text{ at least } N \text{ times before } \hat{\tau}(x, M)) \geq 1 - \varepsilon. \]

**Proof.** To fix our ideas consider the case \(x \geq 0\). Thus \(\hat{\tau}(x, M) = \tau_x + M\).

By ellipticity (condition (B)) it is enough to prove the result with (8) replaced by

\[ \mathbb{P}(X_n \text{ visits } x \text{ at least } N \text{ times before } \tau_{x+M}) \geq 1 - \varepsilon. \]

Let \(\tilde{\tau}_m\) be the first time strictly greater than \(\tau_x\) when either \(|X_{\tilde{\tau}} - x| = m\) or \(X_{\tilde{\tau}} = x\). Pick two numbers \(p, p'\) such that \(\delta L < p' < p < 1\). We claim that if \(m_1\) is large enough then for most environments

\[ (9) \quad \mathbb{P}(X_{\tilde{\tau}_{m_1}} = x) > 1 - p. \]

There are two cases to consider: \(X_{\tau_x + 1} = x + 1\) and \(X_{\tau_x + 1} = x - 1\) (the case \(X_{\tau_x + 1} = x\) is trivial). We consider the first case (the second case is easier).

By Optional Stopping Theorem

\[ \mathbb{P}(X_{\tilde{\tau}_{m_1}} = x + m_1 | X_{\tau_{x+1}} = x+1) = \frac{\mathbb{E}(C_{\tilde{\tau}_{m_1}} - C_{\tau_x}) + 1}{m_1} \leq \frac{Z(x, x + m_1) + 1}{m_1}. \]

So (9) holds if \(Z(x, x + m_1) < m_1 p'\) (observe that we need not impose any restrictions in case \(X_{\tau_x + 1} = x - 1\)). Next

\[ \mathbb{P}(X_{\tilde{\tau}_{m_2}} = x + m_2 | X_{\tilde{\tau}_{m_1}} = x + m_1) = \frac{\mathbb{E}(C_{\tilde{\tau}_{m_2}} - C_{\tilde{\tau}_{m_1}}) + m_1}{m_2} \leq \frac{Z(x, x + m_2) + m_1}{m_2}. \]

Thus if \(\frac{m_1}{m_2} < \frac{p - p'}{2}\) and \(Z(x, x + m_2) < p' m_2\) then

\[ \mathbb{P}(X_{\tilde{\tau}_{m_2}} = x + m_2 | X_{\tilde{\tau}_{m_1}} = x + m_1) < p. \]

Thus if both \(Z(x, x + m_1) < p' m_1\) and \(Z(x, x + m_2) < p' m_2\) then

\[ \mathbb{P}(X_{\tilde{\tau}_{m_2}} = x + m_2) < p^2. \]
Inductively let $m_k$ be the smallest number such that

$$m_k > \frac{2}{p - p'} m_{k-1}.$$ 

Then on $\bigcap_{j=1}^k \{ Z(x, x + m_j) < p'm_j \}$ we have

$$\mathbb{P}(X_{\tau_{m_k}} = x + m_k) < p^k.$$ 

Thus on this set

$$\mathbb{P}(X \text{ returns to } x \text{ before } \tau_{x+m_k}) \geq 1 - p^k.$$ 

Since the amount of cookies between $x$ and $x + m_j$ only decreases between the returns the same argument shows that

$$\mathbb{P}(X \text{ returns to } x \text{ at least } N \text{ times before } \tau_{x+m_k}) \geq (1 - p^k)^N.$$ 

Choose $k$ so that $(1 - p^k)^N > 1 - \varepsilon$. Let $M = m_k$ and $\Omega_M = \bigcap_{j=1}^k \{ Z(0, m_j) \leq p'm_j \}$. Then the Ergodic Theorem implies that if $m_1$ is large enough then $\mathbb{P}(\Omega_M) \geq 1 - \varepsilon$. \hfill \square

**Lemma 6.** For almost all $\omega$, $\frac{C_n - \alpha c_n}{r_n} \to 0$ in probability.

**Proof.** Let $\varepsilon > 0$. Take $N$ such that

$$\sum_{j=N+1}^{\infty} \mathbb{E}(\delta(y, j)) < \varepsilon.$$ 

Split $C_n = C_n^- + C_n^+$, where

$$C_n^- = \sum_k \tilde{\Delta}_k I(l_k \leq N), \quad C_n^+ = \sum_k \tilde{\Delta}_k I(l_k > N).$$

By ergodicity we have $C_n^+ \leq 2\varepsilon r_n$ for large $n$ so the main contribution comes from $C_n^-$. Next

$$C_n^- = \sum_{j=1}^N \sum_{j=1}^* \delta(y, j) I(Q(y, j, n))$$

where $Q(y, j, n)$ is the event that $Y$ visits $y$ at least $j$ times before time $n$ and the meaning of $\sum_{j=1}^*$ is the same as in (5). Take a large number $M$ (the precise conditions on $M$ will be given in equations (14) and (16) below) and split $C_n^- = C_n^0 + C_n^i$ where $C_n^0$ contains the terms $y = (x, s)$ where $x$ is within distance $M$ from either maximum or minimum of $X_k, k \leq n$ and $C_n^i$ contains the remaining terms. Then $C_n^0 \leq 2LMN$ since there are $2LM$ sites within distance $M$ from either maximum or
minimum of $X_k, k \leq n$ and for each site only the first $N$ visits give a non-zero contribution to $C_n^-$. On the other hand

$$C_n^i = \sum_{j=1}^{**} \sum_{j=1}^{N} \delta(y, j) - \sum_{j=1}^{**} \sum_{j=1}^{N} \delta(y, j)I(Q^c(y, j, n))$$

where the summation in (**) runs over $y$ with

$$\min_{k\leq n}(X_k) + M \leq x(y) \leq \max_{k\leq n}(X_k) - M$$

Due to ergodicity for large $n$

$$\left| \sum_{j=1}^{**} \sum_{j=1}^{N} \delta(y, j) - [L \sum_{j=1}^{N} E(\delta(y, j))] \right| \leq \varepsilon r_n$$

and by the choice of $N, L \sum_{j=1}^{N} E(\delta(y, j))$ within $\varepsilon$ from $\delta$. The second term in (10) is less than

$$\hat{C}_n = \sum_{y} \sum_{j=1}^{N} I(\hat{Q}(y, j))$$

where $\hat{Q}((x, s), j, M)$ is the event that the $j$-th visit to $(x, s)$ occurs after time $\hat{\tau}(x, M)$. Therefore to complete the proof of Lemma 6 it remains to show that for almost every $\omega$ given $\varepsilon$ there exists $M$ such that for large $n$ we have

$$P(\hat{C}_n > \varepsilon r_n) < \varepsilon.$$ 

To this end we show that there exists $\eta$ such that

$$P(r_n < \eta \sqrt{n}) < \frac{\varepsilon}{3}.$$ 

Indeed $X_n = B_n + C_n$ and by the Ergodic Theorem for almost every $\omega$ there is a constant $K(\omega)$ such that for all $n$ we have

$$0 < C_n < r_n + K(\omega).$$

Since we also have $|X_n| \leq r_n$ the inequality $r_n < \eta \sqrt{n}$ implies that $B_n < 2\eta \sqrt{n} + K(\omega)$ but by Lemma 3 $P(B_n < 2\eta \sqrt{n} + K(\omega))$ can be made as small as we wish by taking $\eta$ small. Next, by Lemmas 1 and 2

$$P(r_N > R \sqrt{n}) < \frac{\varepsilon}{3}$$

in $R, n$ are sufficiently large. Combining (12) and (13) we get

$$P\left( \frac{\hat{C}_n}{r_n} \leq \frac{\sum_{|x(y)|< R \sqrt{n} \sum_{j=1}^{N} I(\hat{Q}(y, j, M))}{\eta \sqrt{n}} \right) < \frac{2\varepsilon}{3}.$$
Observe that by Lemma 5 we can choose $M$ so large that
\begin{equation}
\Pr(\hat{Q}((x, s), j, M)) \leq \frac{\varepsilon^2 \eta}{100 RN L} + I(\Omega^c_M(T^x, \omega)).
\end{equation}
Therefore
\begin{equation}
\mathbb{E}(\sum_{j=1}^{N} I(\hat{Q}(y, jM))) \leq \frac{\varepsilon^2 \eta \sqrt{n}}{50} + LN \sum_{|x|<R \sqrt{n}} I(\Omega^c_M(T^x, \omega)).
\end{equation}
By Lemma 5 we can take $M$ so large that
\begin{equation}
\Pr(\Omega^c_M) \leq \frac{\varepsilon^2 \eta}{200 RN}.
\end{equation}
Then by ergodicity the last term in (15) is less than $\frac{\varepsilon^2 \eta \sqrt{n}}{50}$ provided that $n$ is sufficiently large. Hence
\begin{equation*}
\mathbb{E} \left( \sum_{j=1}^{N} I(\hat{Q}(y, jM)) \right) \leq \frac{\varepsilon^2 \eta \sqrt{n}}{25}.
\end{equation*}
Therefore by Markov inequality
\begin{equation*}
\Pr \left( \sum_{j=1}^{N} I(\hat{Q}(y, jM)) > \varepsilon \eta \sqrt{n} \right) < \frac{\varepsilon}{25}.
\end{equation*}
This completes the proof of (11). Lemma 6 follows. \hfill \Box

**Proof of Theorem 1.** We have
\begin{equation}
W_n(t) = B_n(t) + C_n(t)
\end{equation}
where $B_n(t)$ and $C_n(t)$ are rescaled versions of the martingale and compensator parts of $X_n$ respectively. By Lemma 4 $W_n(t)$ is tight, by Lemma 3 $B_n(t)$ is tight. Since $C_n(t)$ is a difference of two tight processes it is tight. Accordingly the triple $(W_n(t), B_n(t), C_n(t))$ considered as a family of $\mathbb{R}^3$ valued processes is tight. Let $(W(t), B(t), C(t))$ denote a weak limit of $(W_n(t), B_n(t), C_n(t))$.

By Lemma 3 $B(t) = B(t)$. Hence by (17) we have
\begin{equation}
W(t) = B(t) + C(t).
\end{equation}
Therefore it remains to show that
\begin{equation}
C(t) = \alpha \left[ \max_{[0,t]} W(s) - \min_{[0,t]} W(s) \right]
\end{equation}
since this implies that $W(t)$ satisfies (1) and we will be done by [5].

Hence given $\varepsilon > 0$ there exists $N$ such that
\begin{equation*}
\Pr \left( \max_{|t_2-t_1|<1/N} |C_n(t_2) - C_n(t_1)| \geq \varepsilon \right) \leq \varepsilon.
\end{equation*}
Consequently to establish 18 it is enough to show that for each \( N, \varepsilon \)
\[
\mathbb{P} \left( \exists j < NT \text{ such that } \left| C_n \left( \frac{j}{N} \right) - \alpha \left[ \max_{[0,j/N]} W_n(s) - \min_{[0,j/N]} W_n(s) \right] \right| > \varepsilon \right) \to 0.
\]
Before rescaling this amounts to showing that
\[
\mathbb{P} \left( \left| C_{m_j} - \alpha r_{m_j} \right| \leq \varepsilon \sqrt{n} \text{ for } j = 1 \ldots N \right) \to 1
\]
where \( m_j = nj/N \). Notice that \( r_{m_j} \leq r_n \) and by Lemmas 1 and 2 \( \mathbb{P}(r_n \geq R\sqrt{n}) \) can be made as small as we wish by choosing \( R \) and \( n \) large. Hence it suffices to check that
\[
\mathbb{P} \left( \left| C_{m_j} - \alpha r_{m_j} \right| \leq \varepsilon r_{m_j} \text{ for } j = 1 \ldots N \right) \to 1.
\]
However for fixed \( N \), \( m_j \) runs over a set of finite cardinality \( N \) and so (19) follows from Lemma 6. This concludes the proof of (18). Theorem 1 is established. \( \square \)

References

[8] Kosygina E. and Zerner M. P. W. Positively and negatively excited random walks on integers, with branching processes,