Electron. Commun. Probab. **19** (2014), no. 70, 1–9.
DOI: 10.1214/ECP.v19-3587
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Conditional persistence of Gaussian random walks*

Fuchang Gao† Zhenxia Liu‡ Xiangfeng Yang§

Abstract

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. standard Gaussian random variables, let $S_n = \sum_{i=1}^n X_i$ be the Gaussian random walk, and let $T_n = \sum_{i=1}^n S_i$ be the integrated (or iterated) Gaussian random walk. In this paper we derive the following upper and lower bounds for the conditional persistence:

$$\mathbb{P}\left\{\max_{1\leq k\leq n} T_k \leq 0 \;\middle|\; T_n = 0, S_n = 0\right\} \lesssim n^{-1/2},$$
$$\mathbb{P}\left\{\max_{1\leq k\leq 2n} T_k \leq 0 \;\middle|\; T_{2n} = 0, S_{2n} = 0\right\} \gtrsim \frac{n^{-1/2}}{\log n},$$

for $n \to \infty$, which partially proves a conjecture by Caravenna and Deuschel [3].

Keywords: conditional persistence; random walk; integrated random walk.
AMS MSC 2010: 60G50; 60F99.
Submitted to ECP on June 7, 2014, final version accepted on October 9, 2014.

1 Introduction

Suppose that $X_n, n \geq 1$, are i.i.d. random variables with mean zero and finite positive variance. Denote $S_n = X_1 + X_2 + \cdots + X_n$ and $T_n = S_1 + S_2 + \cdots + S_n$, $n \geq 1$. In this paper, we study the following conjecture of Caravenna and Deuschel [3] which is motivated from their study of sticky particles in a random polymer:

Conjecture: $\mathbb{P}\left\{\max_{1\leq k\leq n} T_k \leq 0 \;\middle|\; T_n = 0, S_n = 0\right\} \asymp n^{-1/2}$.

Here and throughout this paper, the following symbols are used for positive sequences $\alpha(n)$ and $\beta(n)$: $\alpha(n) \lesssim \beta(n)$ if $\limsup_{n\to\infty} \alpha(n)/\beta(n) \leq c_1 < \infty$; $\alpha(n) \gtrsim \beta(n)$ if $\liminf_{n\to\infty} \alpha(n)/\beta(n) \geq c_2 > 0$, where c_1 and c_2 are two positive constants. Furthermore, we denote $\alpha(n) \asymp \beta(n)$ if $\alpha(n) \lesssim \beta(n)$ and $\alpha(n) \gtrsim \beta(n)$. We refer to [3] for the significance of the conjecture and its application in wetting and pinning models. Here we remark that the question is indeed quite natural, by presenting a practical example. Suppose that a person holds n units of shares of a certain stock, of which

*Partially supported by a grant from the Simons Foundation #246211.
†Department of Mathematics, University of Idaho, 83844 Moscow, USA.
E-mail: fuchang@uidaho.edu
‡Blåeldsvägen 12B, 59054 Sturefors, Sweden.
E-mail: zhenxia.liu@hotmail.com
§Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden.
E-mail: xiangfeng.yang@liu.se

the price is assumed to be a general symmetric random walk. The person has two options to sell the stock: either he sells all the n units of shares to get cash now, or he sells one unit of share per period for n periods. If the average rate of increase of the stock price during the n periods is the same as the constant simple interest rate r, and these two options make no difference at the end, then what is the probability that the person never regrets during the n periods after choosing the first option? By the assumptions, the stock price in the period k is $P_k = P_0 + S_k + kr$, where P_0 is the current stock price and $S_k = X_1 + X_2 + \ldots + X_k$ is the random price after k periods with $\{X_n\}_{n\geq 1}$ being i.i.d. symmetric random variables. The person would not regret in the period k if $P_1 + P_2 + \ldots + P_k \leq (P_0 + kr) + (P_0 + (k-1)r) + \ldots + (P_0 + r)$, that is $T_k := S_1 + S_2 + \ldots + S_k \leq 0$. Since there is no difference between the two options after n periods, we have $S_1 + S_2 + \ldots + S_n = 0$. Furthermore, the average rate of increase of the stock price during the n periods is the same as the constant simple interest rate r, therefore $S_n = 0$. Thus, the conditional probability that the person never regrets during the n periods can be expressed exactly as $\mathbb{P}\left\{\max_{1\leq k\leq n} T_k \leq 0 \mid T_n = 0, S_n = 0\right\}$.

The conjecture is quite challenging. In their original paper [3], Caravenna and Deuschel showed that $n^{-11/2} \lesssim \mathbb{P}\left\{\max_{1\leq k\leq n} T_k \leq 0 \mid T_n = 0, S_n = 0\right\} \lesssim (\log n)^{-\alpha}$ for some positive α under a mild assumption on $\{X_n\}$. Recently Aurzuda, Dereich and Lifshits [1] proved that the conjecture holds for the case when $\{X_n\}$ are i.i.d. Bernoulli random variables. Then, Denisov and Wachtel [6] announced an extension of the main result in [1], whose formal proof was not given but claimed to follow from the arguments in [5]. While we believe that the methods proposed in [1] and in [6] for discrete random variables $\{X_n\}$ may be adapted with some appropriate modifications to handle continuous random variables, in this paper we use a more elementary method to study this conjecture for the case when $\{X_n\}$ are i.i.d. standard Gaussian random variables. More precisely, we will prove the following:

Theorem 1.1. *If $\{X_n\}_{n\geq 1}$ are i.i.d. standard Gaussian random variables, $S_n = \sum_{i=1}^n X_i$ and $T_n = \sum_{i=1}^n S_i$, then the following estimates hold*

$$\mathbb{P}\left\{\max_{1\leq k\leq n} T_k \leq 0 \;\middle|\; T_n = 0, S_n = 0\right\} \lesssim n^{-1/2},$$

$$\mathbb{P}\left\{\max_{1\leq k\leq 2n} T_k \leq 0 \;\middle|\; T_{2n} = 0, S_{2n} = 0\right\} \gtrsim \frac{n^{-1/2}}{\log n}$$

as $n \to \infty$.

The main idea of our approach is to write the conditional probability as a ratio of two expectations. For the proof of the upper bound, we write the conditional probability as a ratio of expectations by singling out the middle two random variables $X_{\lfloor n/2\rfloor}$ and $X_{\lfloor n/2\rfloor+1}$, and then reduce the problem to the product of two unconditional persistence probabilities $\mathbb{P}\left\{\max_{1\leq k\leq \lfloor n/4\rfloor} T_k \leq 0\right\}$ and $\mathbb{P}\left\{\max_{\lfloor 3n/4\rfloor\leq k\leq n} \widetilde{T}_k \leq 0\right\}$ (where $\widetilde{T}$ is defined similarly as T using random variables $\{X_k\}_{k\geq\lfloor 3n/4\rfloor}$ instead of $\{X_k\}_{1\leq k\leq\lfloor n/4\rfloor}$). Since both unconditional persistence probabilities are of order $n^{-1/4}$ (cf. [4]; see also [8], [2] and reference therein for other related persistence), the original conditional persistence is of order $n^{-1/2}$. This method works for any continuous random variables $\{X_n\}$ satisfying the corresponding inequality (3.4). For the proof of the lower bound, we rewrite the conditional probability as a ratio of expectations using the last two random variables X_{2n-1} and X_{2n}. Then by the symmetry between the first $n-1$ random variables $X_1, \ldots, X_{n-1}$ and the last $n-1$ random variables $X_n, \ldots, X_{2n-2}$, we arrive at $n^{-1/2}/\log n$. This proof can be also extended to some other random variables (such as exponential random variables) by using central limit theorem. However, a new method seems to be needed to remove the $\log n$ factor.

2 Preparation

For convenience, we introduce some notations. We set

$$S_{k,m} = \begin{cases} X_k + X_{k+1} + \ldots + X_m & \text{if } k \le m \\ X_k + X_{k-1} + \ldots + X_m & \text{if } k > m \end{cases}.$$

Similarly, we denote

$$T_{k,m} = \begin{cases} X_m + 2X_{m-1} + \ldots + (m-k+1)X_k & \text{if } k \le m \\ X_m + 2X_{m+1} + \ldots + (k-m+1)X_k & \text{if } k > m \end{cases}.$$

Thus, $S_{1,m} = S_m$ and $T_{1,m} = T_m$. With these notations, we now can write for $n \geq 4$ and $k+3 < n$,

$$\begin{aligned} S_{1,n} &= S_{1,k} + X_{k+1} + X_{k+2} + S_{n,k+3}, \\ T_{1,n} &= T_{1,k} + (n-k)S_{1,n} - T_{n,k+2}. \end{aligned}$$

Therefore, under the conditions $T_{1,n} = 0$ and $S_{1,n} = 0$, we have

$$\begin{aligned} &S_{1,k} + X_{k+1} + X_{k+2} + S_{n,k+3} = 0, \\ &T_{1,k} - T_{n,k+2} = 0. \end{aligned}$$

Together with the fact that $T_{n,k+2} = T_{n,k+3} + S_{n,k+3} + X_{k+2}$, we obtain

$$\begin{aligned} X_{k+1} &= T_{n,k+3} - T_{1,k} - S_{1,k} := Y_{n-k-2,k}, \\ X_{k+2} &= T_{1,k} - T_{n,k+3} - S_{n,k+3} := Z_{n-k-2,k}. \end{aligned}$$

Furthermore, under the conditions $T_{1,n} = 0$ and $S_{1,n} = 0$,

$$\left\{ \max_{1 \le i \le n} T_{1,i} \le 0 \right\} = \left\{ \max_{1 \le i \le k} T_{1,i} \le 0 \right\} \cap \left\{ \max_{k+3 \le i \le n} T_{n,i} \le 0 \right\}.$$

If we denote $A_m = \{\max_{1 \le i \le m} T_{1,i} \le 0\}$ and $B_m = \{\max_{n-m+1 \le i \le n} T_{n,i} \le 0\}$, then it is straightforward to deduce that

$$\begin{aligned} &\left\{ \max_{1 \le i \le n} T_{1,i} \le 0, S_{1,n} = 0, T_{1,n} = 0 \right\} \\ &= \left\{ \max_{1 \le i \le k} T_{1,i} \le 0, \max_{k+3 \le i \le n} T_{n,i} \le 0, X_{k+1} = Y_{n-k-2,k}, X_{k+2} = Z_{n-k-2,k} \right\} \\ &= A_k \cap B_{n-k-2} \cap \{X_{k+1} = Y_{n-k-2,k}, X_{k+2} = Z_{n-k-2,k}\}. \end{aligned}$$

From the fact that $\{S_{1,n} = 0, T_{1,n} = 0\} = \{X_{k+1} = Y_{n-k-2,k}, X_{k+2} = Z_{n-k-2,k}\}$, it follows

$$\begin{aligned} &\mathbb{P}\left\{ \max_{1 \le i \le n} T_{1,i} \le 0 \;\middle|\; T_{1,n} = 0, S_{1,n} = 0 \right\} \\ &= \mathbb{P}\left\{ A_k \cap B_{n-k-2} \;\middle|\; X_{k+1} = Y_{n-k-2,k}, X_{k+2} = Z_{n-k-2,k} \right\}. \end{aligned}$$

If the density function of X_1 is denoted as $f(x) = (2\pi)^{-1/2} e^{-x^2/2}$, then we claim that

$$q_n := \mathbb{P}\left\{ \max_{1 \le i \le n} T_{1,i} \le 0 \;\middle|\; T_{1,n} = 0, S_{1,n} = 0 \right\} = \frac{\mathbb{E} f(Y_{n-k-2,k}) f(Z_{n-k-2,k}) 1_{A_k} 1_{B_{n-k-2}}}{\mathbb{E} f(Y_{n-k-2,k}) f(Z_{n-k-2,k})}. \tag{2.1}$$

Proof of (2.1). Before the formal proof of (2.1), let us first show an equality which gives a good motivation of (2.1). Suppose that two random variables X and Y are standard Gaussian random variables, and h is a differentiable function, then we will show

$$\mathbb{P}\left\{X \in A \;\middle|\; Y = h(X)\right\} = \frac{\int_A f(x)f(h(x))dx}{\int_{\mathbb{R}} f(x)f(h(x))dx} = \frac{\mathbb{E}f(h(X))1_{\{X\in A\}}}{\mathbb{E}f(h(X))} \tag{2.2}$$

where f is the density function of a standard Gaussian random variable. We can regard (2.2) as the simplest case of (2.1), and these two proofs are essentially the same. The second equality in (2.2) is trivial, so we now prove the first equality in (2.2). A version of the conditional probability can be written as (cf. Section 2.13 in [7])

$$\mathbb{P}\left\{X \in A \;\middle|\; Y = h(X)\right\} = \mathbb{P}\left\{X \in A \;\middle|\; Y - h(X) = 0\right\} = \frac{\int_A f_{X,Y-h(X)}(x,0)dx}{\int_{\mathbb{R}} f_{X,Y-h(X)}(x,0)dx}$$

where $f_{X,Y-h(X)}(\cdot,\cdot)$ denotes the joint density function of the two-dimensional random variable $(X, Y - h(X))$. For notational simplicity, if we let $Z = Y - h(X)$, then the joint density $f_{X,Z}(x,z)$ can be obtained by change of variables from (X,Y) to (X,Z). More precisely, the Jacobian determinant is equal to 1 and $f_{X,Z}(x,z) = f_{X,Y}(x, z+h(x)) = f(x)f(z+h(x))$. Therefore $f_{X,Y-h(X)}(x,0) = f(x)f(h(x))$, which proves (2.2).

Now we come to the proof of (2.1). If we denote $W = (X_1,\dots,X_k,X_{k+3},\dots,X_n)$, then, we can write $Y_{n-k-2,k} = u(W)$ and $Z_{n-k-2,k} = v(W)$ where u, v are functions on $\mathbb{R}^{n-2}$. Let g be the density function of W. Because W and X_{k+1} and X_{k+2} are independent, the joint density of W, X_{k+1} and X_{k+2} is $g(w)f(x_{k+1})f(x_{k+2})$. Thus, as in (2.2), the conditional density of $(W \mid X_{k+1} = u(W), X_{k+2} = v(W))$ could be given as

$$\frac{g(w)f(u(w))f(v(w))}{\int_{\mathbb{R}^{n-2}} g(w)f(u(w))f(v(w))dw}.$$

Since $u(W) = Y_{n-k-2,k}$ and $v(W) = Z_{n-k-2,k}$, the denominator can be written as

$$\int_{\mathbb{R}^{n-2}} g(w)f(u(w))f(v(w))dw = \mathbb{E}f(u(W))f(v(W)) = \mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k}).$$

Therefore,

$$\begin{aligned}
q_n &= \mathbb{P}\left\{A_k \cap B_{n-k-2} \;\middle|\; X_{k+1} = Y_{n-k-2,k}, X_{k+2} = Z_{n-k-2,k}\right\} \\
&= \int_{a_k \cap b_{n-k-2}} \frac{g(w)f(u(w))f(v(w))}{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})} dw \\
&= \frac{\mathbb{E}f(u(W))f(v(W))1_{A_k \cap B_{n-k-2}}}{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})} \\
&= \frac{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})1_{A_k}1_{B_{n-k-2}}}{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})},
\end{aligned}$$

where $a_m = \{\max_{1\le i\le m} t_{1,i} \le 0\}$, $b_m = \{\max_{n-m+1\le i\le n} t_{n,i} \le 0\}$, $s_{k,m}$ and $t_{k,m}$ are defined similarly as $S_{k,m}$ and $T_{k,m}$ with $\{X_i\}$ replaced by $\{x_i\}$. □

3 Upper Bound

To prove the upper bound, we choose $k = \lfloor n/2 \rfloor - 1$ and $m = \lfloor k/2 \rfloor$. Because $A_k \subseteq A_m$ and $B_{n-k-2} \subseteq B_m$, it follows from (2.1) that

$$q_n \le \frac{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})1_{A_m}1_{B_m}}{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})}. \tag{3.1}$$

We now take a closer look at $Y_{n-k-2,k}$ and $Z_{n-k-2,k}$. For $k+3+m<n$, we can write

$$\begin{aligned}
Y_{n-k-2,k} =& T_{n,k+3} - T_{1,k} - S_{1,k}\\
=& [T_{n,k+3+m} + mS_{n,k+3+m} - T_{1,k-m} - (m+1)S_{1,m}]\\
&+ [T_{k+2+m,k+3} - T_{k-m+1,k} - S_{k-m+1,k}]\\
:=& a + U,
\end{aligned}$$

and

$$\begin{aligned}
Z_{n-k-2,k} =& T_{1,k} - T_{n,k+3} - S_{n,k+3}\\
=& [T_{1,k-m} + mS_{1,k-m} - T_{k,k+3+m} - (m+1)S_{n,k+3+m}]\\
&+ [T_{k-m+1,k} - T_{k+2+m,k+3} - S_{k+2+m,k+3}]\\
:=& b + V.
\end{aligned}$$

With these notations, (3.1) can be rewritten as

$$q_n \leq \frac{\mathbb{E}f(a+U)f(b+V)1_{A_m}1_{B_m}}{\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})}. \tag{3.2}$$

Note that a, b, 1_{A_m} and 1_{B_m} only depend on $X_1,...,X_{k-m},X_{k+m+3},...,X_n$, while U and V only depend on $X_{k-m+1},...,X_k,X_{k+3},...,X_{k+m+2}$. Therefore, a, b, 1_{A_m} and 1_{B_m} are independent of (U,V). If we can show that there exists a constant $C>0$ such that for all real numbers α and β,

$$\mathbb{E}f(\alpha+U)f(\beta+V) \leq C\cdot\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k}), \tag{3.3}$$

then by conditioning on the variables $X_1,...,X_{k-m},X_{k+m+3},...,X_n$, we can bound the numerator on the right-hand side of (3.2) by $C\cdot\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k})\cdot\mathbb{E}(1_{A_m}1_{B_m})$. Thus, we immediately obtain $q_n\leq C\cdot\mathbb{P}\{A_m\cap B_m\}$. By the unconditional persistence estimate obtained in [4], we have $\mathbb{P}\{A_m\}=\mathbb{P}\{B_m\}\leq C'm^{-1/4}$. Thus $q_n\leq C''n^{-1/2}$.

Note that (U,V) has the same distribution as $(Y_{m,m},Z_{m,m})$. Thus (3.3) is equivalent to the following claim: there exists a constant C such that for all real number α and β,

$$\mathbb{E}f(\alpha+Y_{m,m})f(\beta+Z_{m,m}) \leq C\cdot\mathbb{E}f(Y_{n-k-2,k})f(Z_{n-k-2,k}) \tag{3.4}$$

for $n\geq 4$, $k=\lfloor n/2\rfloor-1$ and $m=\lfloor k/2\rfloor$.

It remains to show the claim. To this end, we prove the following lemma.

Lemma 3.1. *If U and V are two centered Gaussian random variables, then for any $\alpha,\beta\in\mathbb{R}$,*

$$\mathbb{E}e^{-\frac{1}{2}(U+\alpha)^2}e^{-\frac{1}{2}(V+\beta)^2} = \frac{1}{\sigma}\exp\left\{-\frac{(1+\mathbb{E}V^2)\alpha^2+(1+\mathbb{E}U^2)\beta^2-2\alpha\beta\mathbb{E}UV}{2\sigma^2}\right\}$$

where $\sigma^2=(1+\mathbb{E}U^2)(1+\mathbb{E}V^2)-(\mathbb{E}UV)^2$.

Proof. Without loss of generality, we can assume $U=\sigma_U X$, and $V=\sigma_V(\rho X+\sqrt{1-\rho^2}Y)$, where X and Y are independent $N(0,1)$ random variables, and $\rho=\mathrm{corr}(U,V)$. Conditioning on X and using the identity

$$\mathbb{E}e^{-\frac{1}{2}(cY+t)^2} = \frac{1}{\sqrt{1+c^2}}e^{-\frac{t^2}{2(1+c^2)}} \tag{3.5}$$

which holds for all $c, t \in \mathbb{R}$, we obtain

$$\begin{aligned}\mathbb{E}\left[e^{-\frac{1}{2}(U+\alpha)^2}e^{-\frac{1}{2}(V+\beta)^2} \mid X\right] &= \frac{1}{\sqrt{1+\sigma_V^2(1-\rho^2)}}e^{-\frac{1}{2}(\sigma_U X+\alpha)^2-\frac{1}{2}\frac{(\sigma_V\rho X+\beta)^2}{1+\sigma_V^2(1-\rho^2)}} \\ &:= \frac{1}{\sqrt{1+\sigma_V^2(1-\rho^2)}}e^{-\frac{1}{2}(AX+B)^2-\frac{1}{2}C},\end{aligned}$$

where

$$\begin{aligned}A &= \sqrt{\sigma_U^2+\frac{\sigma_V^2\rho^2}{1+\sigma_V^2(1-\rho^2)}}, \\ B &= \frac{1}{A}\left(\sigma_U\alpha+\frac{\sigma_V\rho\beta}{1+\sigma_V^2(1-\rho^2)}\right), \\ C &= \alpha^2+\frac{\beta^2}{1+\sigma_V^2(1-\rho^2)}-B^2.\end{aligned}$$

Taking expectation and using (3.5) again, we obtain

$$\mathbb{E}e^{-\frac{1}{2}(U+\alpha)^2}e^{-\frac{1}{2}(V+\beta)^2} = \frac{1}{\sqrt{[1+\sigma_V^2(1-\rho^2)](1+A^2)}}e^{-\frac{B^2}{2(1+A^2)}-\frac{C}{2}},$$

which proves the lemma after simplification. □

Note that for all $\alpha, \beta \in \mathbb{R}$,

$$\exp\left\{-\frac{(1+\mathbb{E}V^2)\alpha^2+(1+\mathbb{E}U^2)\beta^2-2\alpha\beta\mathbb{E}UV}{2\sigma^2}\right\} \le e^{-\frac{\alpha^2+\beta^2}{2\sigma^2}}.$$

The lemma above applied twice implies the following inequality:

$$\mathbb{E}e^{-(\alpha+U)^2/2}e^{-(\beta+V)^2/2} \le e^{-\frac{\alpha^2+\beta^2}{2\sigma^2}}\mathbb{E}e^{-U^2/2}e^{-V^2/2}. \tag{3.6}$$

With $a, b, Y_{m,m}$ and $Z_{m,m}$ defined between (3.1) and (3.4), by applying (3.6) followed by Lemma 3.1 for $\alpha = \beta = 0$, we obtain

$$\begin{aligned}\mathbb{E}e^{-(a+Y_{m,m}^2)/2}e^{-(b+Z_{m,m}^2)/2} &\le \mathbb{E}e^{-Y_{m,m}^2/2}e^{-Z_{m,m}^2/2} \\ &= [(1+\mathbb{E}|Y_{m,m}|^2)(1+\mathbb{E}|Z_{m,m}|^2)-(\mathbb{E}Y_{m,m}Z_{m,m})^2]^{-1/2} \\ &= \frac{\sqrt{3}}{(m+1)\sqrt{(2m+1)(2m+3)}}.\end{aligned}$$

Similarly, for $k = \lfloor n/2 \rfloor - 1$ defined above, if n is even, then $n = 2k+2$, we have

$$\mathbb{E}e^{-\frac{1}{2}Y_{n-k-2,k}^2-\frac{1}{2}Z_{n-k-2,k}^2} = \frac{\sqrt{3}}{(k+1)\sqrt{(2k+1)(2k+3)}};$$

if n is odd, we have $n = 2k+3$, and

$$\mathbb{E}e^{-\frac{1}{2}Y_{n-k-2,k}^2-\frac{1}{2}Z_{n-k-2,k}^2} = \frac{\sqrt{6}}{2\sqrt{(k+1)(k+2)(2k+3)}}.$$

In either case, since $m = \lfloor k/2 \rfloor$, we immediately obtain (3.4) for $C \approx \sqrt{8}$. This finishes the proof of the upper bound.

4 Lower Bound

The idea of the proof of the lower bound is similar to that of the upper bound. We first introduce a few more notations. For a fixed large n, we define two functions F_1 and F_2 as

$$\begin{aligned} F_1(y_1,\dots,y_n) &= f(y_1)f(-2y_1+y_2)f(y_1-2y_2+y_3)\dots f(y_{n-2}-2y_{n-1}+y_n), \\ F_2(y_{n+3},\dots,y_{2n+2}) &= f(y_{n+3}-2y_{n+4}+y_{n+5})\dots f(y_{2n}-2y_{2n+1}+y_{2n+2}) \\ &\qquad \cdot f(y_{2n+1}-2y_{2n+2})f(y_{2n+2}), \end{aligned}$$

and four sets

$$\begin{aligned} \Omega^+ &= \left\{(y_1,\dots,y_{2n+2}) \in \mathbb{R}^{2n+2} : \min_{1\le k\le 2n+2} y_k \ge 0\right\}, \\ \Omega_1^+ &= \left\{(y_1,\dots,y_n) \in \mathbb{R}^n : \min_{1\le k\le n} y_k \ge 0\right\}, \\ \Omega_2^+ &= \left\{(y_{n+3},\dots,y_{2n+2}) \in \mathbb{R}^n : \min_{n+3\le k\le 2n+2} y_k \ge 0\right\}, \\ \Omega_3^+ &= \{y_{n+1}\ge 0, y_{n+2}\ge 0\}. \end{aligned}$$

For notational simplicity, we will derive a lower bound for q_{2n+4} instead of q_{2n}. This of course makes no essential difference. Note that

$$\begin{aligned} q_{2n+4} &= \mathbb{P}\left\{\max_{1\le k\le 2n+4} T_k \le 0 \;\middle|\; T_{2n+4}=0, S_{2n+4}=0\right\} \\ &= \mathbb{P}\left\{\min_{1\le k\le 2n+4} T_k \ge 0 \;\middle|\; T_{2n+4}=0, S_{2n+4}=0\right\} \\ &= \frac{\mathbb{E}\left[e^{-T_{2n+2}^2/2}e^{-(T_{2n+2}+S_{2n+2})^2/2}1_{\{\min_{1\le k\le 2n+2} T_k\ge 0\}}\right]}{\mathbb{E}\left[e^{-T_{2n+2}^2/2}e^{-(T_{2n+2}+S_{2n+2})^2/2}\right]}. \end{aligned} \tag{4.1}$$

The denominator can be directly computed using Lemma 3.1:

$$\mathbb{E}\left[e^{-T_{2n+2}^2/2}e^{-(T_{2n+2}+S_{2n+2})^2/2}\right] = \frac{1}{(2n+4)\sqrt{\frac{(2n+3)(2n+5)}{12}}} \asymp n^{-2}.$$

We thus focus on the numerator

$$\mathbb{E}\left[e^{-T_{2n+2}^2/2}e^{-(T_{2n+2}+S_{2n+2})^2/2}1_{\{\min_{1\le k\le 2n+2} T_k\ge 0\}}\right],$$

which can be expressed as a multiple integral with respect to the joint distribution of $\{X_1,\dots,X_{2n+2}\}$. But here we choose a multiple integral with respect to the joint distribution of $\{T_1,\dots,T_{2n+2}\}$. We do the following change of variables

$$X_1 = T_1, \quad X_2 = T_2 - 2T_1, \quad X_3 = T_3 - 2T_2 + T_1, \quad \dots, \quad X_{2n+2} = T_{2n+2} - 2T_{2n+1} + T_{2n}.$$

It is then straightforward to check that the Jacobian determinant is 1. Thus, the numerator becomes

$$\begin{aligned} &\mathbb{E}\left[e^{-T_{2n+2}^2/2}e^{-(T_{2n+2}+S_{2n+2})^2/2}1_{\{\min_{1\le k\le 2n+2} T_k\ge 0\}}\right] \\ &= \int_{\mathbb{R}^{2n+2}} \frac{1}{\left(\sqrt{2\pi}\right)^{2n+2}} \exp\left\{-\frac{y_1^2}{2} - \frac{(-2y_1+y_2)^2}{2} - \frac{(y_1-2y_2+y_3)^2}{2} - \dots\right. \end{aligned}$$

$$
\begin{aligned}
&\qquad -\frac{(y_{2n}-2y_{2n+1}+y_{2n+2})^2}{2}-\frac{(y_{2n+1}-2y_{2n+2})^2}{2}-\frac{y_{2n+2}^2}{2}\Big\} \\
&\qquad \cdot 1_{\{\min_{1\le k\le 2n+2} y_k\ge 0\}}dy_1\dots dy_{2n+2} \\
&= 2\pi\int_{\Omega^+}F_1(y_1,\dots,y_n)f(y_{n-1}-2y_n+y_{n+1})f(y_n-2y_{n+1}+y_{n+2}) \\
&\qquad f(y_{n+1}-2y_{n+2}+y_{n+3})f(y_{n+2}-2y_{n+3}+y_{n+4})F_2(y_{n+3},\dots,y_{2n+2})dy_1\dots dy_{2n+2} \\
&= 2\pi\int_{\Omega_3^+}\Big\{\int_{\Omega_1^+}F_1(y_1,\dots,y_n)f(y_{n-1}-2y_n+y_{n+1})f(y_n-2y_{n+1}+y_{n+2})dy_1\dots dy_n \\
&\int_{\Omega_2^+}f(y_{n+1}-2y_{n+2}+y_{n+3})f(y_{n+2}-2y_{n+3}+y_{n+4})F_2(y_{n+3},\dots,y_{2n+2})dy_{n+3}\dots dy_{2n+2}\Big\} \\
&\qquad\qquad dy_{n+1}dy_{n+2} \\
&:= 2\pi\int_{\Omega_3^+}G_1(y_{n+1},y_{n+2})G_2(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2} \\
&= 2\pi\int_{\Omega_3^+}G_1^2(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}
\end{aligned}
$$

where the last equality comes from the symmetry of $\{F_i\}_{i=1,2}$ and f.

In order to estimate the last integral, we consider a subset D of Ω_3^+ defined as

$$
\begin{aligned}
D=\Big\{&(y_{n+1},y_{n+2})\in\mathbb{R}^2: y_{n+1}\ge 0, y_{n+2}\ge 0,\ \text{and} \\
&\qquad y_{n+1}<n^{3/2}(\log n)^{1/2}, |y_{n+1}-y_{n+2}|<\sqrt{n}(\log n)^{1/2}\Big\}.
\end{aligned}
$$

The area $|D|$ of the region D is $|D|\asymp n^2\log n$. By applying Hölder's inequality, we obtain

$$
\begin{aligned}
&\int_{\Omega_3^+}G_1^2(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2} \\
&\ge \frac{1}{|D|}\left(\int_D G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}\right)^2 \\
&= \frac{1}{|D|}\left(\int_{\Omega_3^+}G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}-\int_{\Omega_3^+\setminus D}G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}\right)^2.
\end{aligned}
\tag{4.2}
$$

By definition and using the unconditional persistence probability of [4], the first integral can be estimated as

$$
\int_{\Omega_3^+}G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}=\mathbb{P}\left\{\min_{1\le k\le n+2}T_k\ge 0\right\}\asymp n^{-1/4}. \tag{4.3}
$$

The second integral over $\Omega_3^+\setminus D$ can be estimated as follows. From definition,

$$
\begin{aligned}
&\int_{\Omega_3^+\setminus D}G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2} \\
&= \mathbb{P}\left\{\min_{1\le k\le n+2}T_k\ge 0\ \cap\ \left(|T_{n+1}|>n^{3/2}(\log n)^{1/2}\ \cup\ |T_{n+1}-T_{n+2}|>\sqrt{n}(\log n)^{1/2}\right)\right\} \\
&\le \mathbb{P}\left\{|T_{n+1}|>n^{3/2}(\log n)^{1/2}\right\}+\mathbb{P}\left\{|T_{n+1}-T_{n+2}|>\sqrt{n}(\log n)^{1/2}\right\}.
\end{aligned}
$$

Since T_{n+1} is a Gaussian random variable with mean zero and variance $n^3/3+n^2/2+n/6$,

$$
\mathbb{P}\left\{|T_{n+1}|>n^{3/2}(\log n)^{1/2}\right\}\le\frac{const.}{(\log n)^{1/2}}\exp\left\{-\frac{\log n}{2}\right\}\lesssim n^{-1/2}.
$$

Similarly, we deduce that $\mathbb{P}\left\{|T_{n+1}-T_{n+2}|>\sqrt{n}(\log n)^{1/2}\right\}\lesssim n^{-1/2}$. Therefore,

$$\int_{\Omega_3^+\setminus D} G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}\lesssim n^{-1/2}.$$

Combining this with (4.3), we conclude from (4.2) that

$$\begin{aligned}
&\int_{\Omega_3^+} G_1^2(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}\\
&\geq \frac{1}{|D|}\left(\int_{\Omega_3^+} G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}-\int_{\Omega_3^+\setminus D} G_1(y_{n+1},y_{n+2})dy_{n+1}dy_{n+2}\right)^2\\
&\asymp \frac{1}{|D|}\cdot n^{-1/2}\asymp n^{-5/2}(\log n)^{-1}.
\end{aligned}$$

This, together with the estimate of the denominator in (4.1), yields

$$q_{2n+4}\gtrsim \frac{1}{n^{1/2}\log n},$$

which completes the proof of the lower bound.

References

[1] Aurzada, F., Dereich, S. and Lifshits, M.: Persistence probabilities for an integrated random walk bridge. *Probab. Math. Statist.* **34**, (2014), 1–22.

[2] Aurzada, F. and Simon, T.: Persistence probabilities & exponents, arXiv:1203.6554

[3] Caravenna, F. and Deuschel, J.: Pinning and wetting transition for $(1+1)$-dimensional fields with Laplacian interaction. *Ann. Probab.* **36**, (2008), 2388–2433. MR-2478687

[4] Dembo, A., Ding, J. and Gao, F.: Persistence of iterated partial sums. *Ann. Inst. Henri Poincaré Probab. Stat.* **49**, (2013), 873–884. MR-3112437

[5] Denisov, D. and Wachtel, V.: Random walks in cones, arXiv:1110.1254

[6] Denisov, D. and Wachtel, V.: Exit times for integrated random walks, arXiv:1207.2270

[7] Gut, A.: Probability: a graduate course. Springer, New York, 2013. xxvi+600 pp. MR-2977961

[8] Vysotsky, V.: Positivity of integrated random walks. *Ann. Inst. Henri Poincaré Probab. Stat.* **50**, (2014), 195–213. MR-3161528

Acknowledgments. We are grateful to an anonymous referee and the Editor for valuable comments and suggestions which improved the presentation of the paper. The first named author also thanks Amir Dembo for bringing up the conjecture to his attention, and for several helpful discussions.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])
- Non profit, sponsored by IMS[2], BS[3], PKP[4]
- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]OJS: Open Journal Systems `http://pkp.sfu.ca/ojs/`
[2]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[3]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[4]PK: Public Knowledge Project `http://pkp.sfu.ca/`
[5]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`