Weakly Self-Avoiding Walks on Graphs and Self-Intersection Events

Irene Hueter*

Abstract

We consider the number J_n of self-intersections of a weakly self-avoiding walk with parameter $\beta > 0$ of length n on an infinite, locally finite connected transitive graph of degree ≥ 2. We show the large deviation type result that for every fixed $\beta > 0$ there are positive finite constants $b_1 < b_2$ and τ such that, for all large enough n, the probability of the event $\{J_n \in [b_1 n, b_2 n]\}$ exceeds $1 - e^{-n\tau}$.

1 Introduction

The self-avoiding walk (SAW) kept much of its fascination ever since the chemist Paul Flory [2] called upon this model in 1949 when he observed that the end-to-end distance of a long linear polymer chain in 3 dimensions must have a power of the chain length larger than 1/2 and should approach 0.6. This problem on the displacement exponent of the SAW together with an entire collection of problems is still unsettled. Some of the walks that allow self-intersections but penalize them, for example, the weakly self-avoiding walk, closely mimic the SAW, a symmetric nearest neighbor random walk that visits any site of a lattice or graph no more than once. These self-intersections may be exploited to gain insight into the behavior of the SAW. A number of repelling walks enjoy the same open problems as the SAW and are conjectured to inherit features from the SAW or vice versa, for instance, the values of the critical exponents, including the displacement exponent.

In this paper, our interest is in the concentration of mass of the number of self-intersection events at sites or bonds of the weakly SAW in $\mathbf{Z}^d$ for every $d \geq 1$. We determine an interval which, for a large step size, contains this number with overwhelming probability. More precisely, if J_n and $\mathbf{Q}_n^\beta$ denote the number of self-intersections and the probability measure of the weakly SAW with parameter $\beta > 0$ of length n (defined in (2.1) and (2.3) below), then we prove in Theorem 1 (Section 4) that, for every fixed $\beta > 0$, there are positive finite constants $b_1 < b_2$, and τ (to be detailed further) such that for all sufficiently large n,

$$\mathbf{Q}_n^\beta(J_n \in [b_1 n, b_2 n]) > 1 - \exp\{-n\,\tau\}.$$

*City University of New York, Baruch College. Partially supported by PSC-CUNY grant #60032-33-34.

Date: July 3, 2003

AMS 2000 Mathematics Subject Classification: *Primary:* 60G50. *Secondary:* 60D05, 60G57.

Key words and phrases: path restrictions, self-avoiding trail, self-avoiding walk, self-intersection events, transitive graphs, weakly self-avoiding walks.

There are reasons to believe that this large deviation type property of J_n is helpful in analyzing the displacement exponent of the SAW in $\mathbf{Z}^2$ or $\mathbf{Z}^d$ for $d > 2$ (see Hueter [4]) and of related polymer models. This likely order of J_n, which is much smaller than expected of the symmetric simple random walk in $\mathbf{Z}$ and $\mathbf{Z}^2$, hints at the stronger repulsion than common to the random walk. The exponential decay with n of the tails of the distribution of J_n extends to other walks restricted by counts of patterns in their paths and is not limited to the Euclidean lattices but shared by walks on a large class of lattices and infinite, locally finite, connected graphs, for instance, the transitive graphs, as we will explain.

Self-avoiding walks mathematically model linear polymers as observed in chemistry (Flory [2, 3]), physics, and biological systems. Apart from Monte Carlo simulations and numerical, speculative, and heuristic work – an enormous collection on its own – the field of self-avoiding walks is widely open and lies on almost untouched ground as yet. The interested reader is encouraged to consult Madras and Slade [5] for an extensive survey.

Showing the upper bound on J_n is straightforward (Proposition 1). The ideas of the somewhat involved, though, elementary proof of the lower bound on J_n (Proposition 2) are the following: we pick any path with "comparatively few" self-intersections, introduce to it much more self-intersections via suitably placed "backtracking steps," count these newly generated paths, and verify that the number of the latter paths considerably exceeds the number of the former paths in that it sets off the larger penalizing weight and hence contributes significantly more to the partition function of the weakly SAW. An implication thus is that the latter paths have significantly larger probability than the former paths to occur in the weakly SAW model. The point of the approach is that there are more locations on the path to choose from to place additional self-intersections than locations already occupied by the smaller number of self-intersections that exist. While it is plausible that large deviation techniques as well handle our results, we do not pursue this here.

The outline of the paper is as follows. Section 2 describes the weakly self-avoiding walks and trails. Section 3 reviews background on the paths of the simple random walk without self-intersection events and as well discusses consequences for related models that weigh paths according to counts of geometric restrictions and patterns. Section 4 is devoted to the proof of the main result and ends with a few open problems.

Acknowledgment. I like to thank an anonymous referee for a thorough reading of an earlier version of this paper, for a number of suggestions that helped to improve its exposition and for providing an idea that simplified the proof of Proposition 2.

2 Weakly Self-Avoiding Walks and Trails

Assume that $\mathcal{G}$ is an infinite, locally finite connected transitive graph with degree $\deg(\mathcal{G}) \geq 2$. An infinite graph $\mathcal{G}$ is *locally finite* if each vertex has bounded degree (the degree of a vertex of $\mathcal{G}$ is the number of edges which emanate from that vertex). We say that $\mathcal{G}$ is *transitive* (or homogeneous) if, for any two vertices x and y of $\mathcal{G}$, there is a graph automorphism which maps x to y. Examples of transitive graphs include the integer lattices $\mathbf{Z}^d$, the triangular lattice and the honey-comb lattice in the plane, and homogeneous trees. Let $\mathbf{0}$ denote a distinguished vertex of $\mathcal{G}$ and let $S_0 = \mathbf{0}, S_1, \ldots, S_n$ denote a symmetric simple random walk (SRW) on $\mathcal{G}$ of length n, starting at $\mathbf{0}$. Let Σ_n denote the set of these SRW-paths $\omega = S_0, S_1, \ldots, S_n$ of length n. Define the *number of self-intersections* or the *self-intersection local time* (SILT) at *sites* or vertices of $\mathcal{G}$ by

$$J_n = J_n(S_0, S_1, \ldots, S_n) = \sum_{0 \le i < j \le n} 1_{\{S_i = S_j\}} \tag{2.1}$$

and the SILT at *bonds* or edges of $\mathcal{G}$ by

$$\tilde{J}_n = \tilde{J}_n(S_0, S_1, \ldots, S_n) = \sum_{0 < i < j \le n} 1_{\{S_i - S_{i-1} = \pm(S_j - S_{j-1})\}}, \tag{2.2}$$

where $S_i - S_{i-1} = \pm(S_j - S_{j-1})$ means that the vector $S_i - S_{i-1}$ either equals $S_j - S_{j-1}$ or equals $S_{j-1} - S_j$. Let $|\cdot|$ denote cardinality. If $\beta \ge 0$ denotes the self-intersection parameter, then the *weakly self-avoiding walk* or *Domb-Joyce model* [1] is the stochastic process that is induced by the probability measure

$$\mathbf{Q}_n^\beta(S_0, S_1, \ldots, S_n) = \frac{\exp\{-\beta J_n\}}{|\Sigma_n| \, \mathbf{E} \exp\{-\beta J_n\}} \tag{2.3}$$

on Σ_n, where $\mathbf{E} \exp\{-\beta J_n\} = |\Sigma_n|^{-1} \sum_{\omega \in \Sigma_n} \exp\{-\beta J_n\}$ and $\mathbf{E}$ stands for the expectation of the SRW. In the same fashion, we define the *weakly self-avoiding trail* or the *weakly self-avoiding walk with edge penalization* as the stochastic process that is induced by the probability measure

$$\tilde{\mathbf{Q}}_n^\beta(S_0, S_1, \ldots, S_n) = \frac{\exp\{-\beta \tilde{J}_n\}}{|\Sigma_n| \, \mathbf{E} \exp\{-\beta \tilde{J}_n\}} \tag{2.4}$$

on Σ_n. These measures weigh the simple random walks of length n with the exponential penalizing factor $\exp\{-\beta r\}$, where r represents the number of self-intersection events at sites (vertices of $\mathcal{G}$) or bonds (edges of $\mathcal{G}$), respectively. When setting $\beta = 0$, we gain back the symmetric SRW from both models, while letting $\beta \to \infty$ mimics the *self-avoiding walk* (SAW) on $\mathcal{G}$ or the *self-avoiding trail* (SAT) on $\mathcal{G}$. The former walk visits each site of its path exactly once, whereas the latter walk visits each bond of its path exactly once. Thus, the SAW is a symmetric SRW that avoids to visit any site in $\mathcal{G}$ more than once while the SAT is a symmetric SRW that avoids to visit any bond in $\mathcal{G}$ more than once. Let $\mathbf{E}_\beta = \mathbf{E}_{\mathbf{Q}_n^\beta}$ and $\tilde{\mathbf{E}}_\beta = \mathbf{E}_{\tilde{\mathbf{Q}}_n^\beta}$ denote expectations under the measures $\mathbf{Q}_n^\beta$ and $\tilde{\mathbf{Q}}_n^\beta$, respectively. Note that $\mathbf{E}_0 = \tilde{\mathbf{E}}_0$ denotes expectation relative to the SRW. For every integer $n \ge 1$, the SAW-measure is defined to be the uniform measure on the set of all self-avoiding paths of length n and the SAT-measure is the uniform measure on the set of self-avoiding trails of length n.

On any connected graph, a walk which returns to a bond on its path has to return to at least one site on its path. In the opposite direction, on some graphs, it is possible for the walk to visit some site for a second time without traversing any edge for a second time. We hence see that

$$\tilde{J}_n \le J_n.$$

On a tree, however, it is impossible to visit any vertex for a second time without traversing some edge for a second time, and thus, $J_n = \tilde{J}_n$. In particular, the two measures $\mathbf{Q}_n^\beta$ and $\tilde{\mathbf{Q}}_n^\beta$ coincide on any tree. A moment's thought shows that paths that exhibit $\tilde{J}_n = 0$ and $J_n > 0$ must exhibit *loops*. Of course, $J_n = 0$ prohibits loops. Those loops are genuine loops (and non-collapsed) in the sense that they have interior when drawn in the plane, say, or have space inside them.

3 Walks that Avoid Bonds or Sites of their Paths

This section attempts to explain in which generality self-intersection events of walks on graphs enjoy the property highlighted in this note, reminiscent of large deviation estimates. We first look at the collection $\Gamma_n \subset \Sigma_n$ of SAW in $\mathcal{G}$ of length n that start at $\mathbf{0}$, avoiding the sites of their own paths. Since we assumed that $\mathcal{G}$ is transitive, concatenating two paths in $\mathcal{G}$ that start at $\mathbf{0}$ generates a path that is in $\mathcal{G}$ and begins at $\mathbf{0}$. All paths in $\mathcal{G}$ that begin at $\mathbf{0}$ can be thought of as the concatenation of certain pairs of paths in $\mathcal{G}$. If we concatenate two paths $\gamma_1 \in \Gamma_n$ and $\gamma_2 \in \Gamma_m$ for any pair (m, n) of nonnegative integers, we do not always end up with a path in Γ_{m+n}. Therefore, we find

$$|\Gamma_{n+m}| \leq |\Gamma_n| \cdot |\Gamma_m|, \tag{3.1}$$

where $|\Gamma_n|$ denotes the cardinality of Γ_n. An easy supermultiplicativity argument yields that the limit

$$\lim_{n\to\infty} |\Gamma_n|^{1/n} = e^{\omega(\mathcal{G})} \tag{3.2}$$

exists for some $1 \leq e^{\omega(\mathcal{G})} \leq \deg(\mathcal{G}) - 1$, often called the *connective constant* of the SAW, and that

$$|\Gamma_n| \geq e^{n\omega(\mathcal{G})} \tag{3.3}$$

for every integer $n \geq 0$. The *upper* bound $\deg(\mathcal{G}) - 1$ may be obtained by overestimating $|\Gamma_n|$ and counting all paths of length n which do not return to the most recently visited site of the path. Depending on the particular graph, the bounds on $\omega(\mathcal{G})$ can be improved readily yet, as of now, it is not known how to derive the precise value of $\omega(\mathcal{G})$ for any $\mathcal{G}$ except for $\mathbf{Z}$, of course, and some trees. There is an ample gallery of numerical work on this issue (for an entrance point to the literature, see Madras and Slade [5]). On trees, Γ_n consists of those paths that start at the root vertex and end at one of the vertices at distance n from the root vertex. The cardinality $|\Gamma_n|$ can be determined for certain classes of trees.

We next consider the set $\tilde{\Gamma}_n$ of self-avoiding trails in $\mathcal{G}$ of length n, the paths without self-intersections at bonds. Since a path may avoid its own bonds but exhibit any number of self-intersections at sites, and thus, does not avoid all of the sites visited, we realize that

$$\Gamma_n \subset \tilde{\Gamma}_n \tag{3.4}$$

for every $n \geq 0$, with equality of the sets holding on trees. In fact, the difference

$$\mathcal{L}_n = \tilde{\Gamma}_n - \Gamma_n \tag{3.5}$$

of the two sets consists of *exactly* those paths of length n that have loops. Since the loops may be nested and be arbitrarily large, the set of points at which the loop is pinned to the path, and thus, experience a self-intersection event does not determine the set $\mathcal{L}_n$ in any strong sense. After this digression, let us return to estimate the cardinality $|\tilde{\Gamma}_n|$. Again, because, for any pair (m, n), concatenating two paths $\gamma_1 \in \tilde{\Gamma}_n$ and $\gamma_2 \in \tilde{\Gamma}_m$ does not always produce a path that avoids its own bonds, we obtain

$$|\tilde{\Gamma}_{n+m}| \leq |\tilde{\Gamma}_n| \cdot |\tilde{\Gamma}_m|. \tag{3.6}$$

In view of a supermultiplicativity argument as before, the limit

$$\lim_{n\to\infty} |\tilde{\Gamma}_n|^{1/n} = e^{\tilde{\omega}(\mathcal{G})} \tag{3.7}$$

exists for some $e^{\omega}(\mathcal{G}) \leq e^{\tilde{\omega}(\mathcal{G})} \leq \deg(\mathcal{G}) - 1$ and $|\tilde{\Gamma}_n| \geq e^{n\tilde{\omega}(\mathcal{G})}$ for every integer $n \geq 0$. As a consequence of (3.4),

$$\omega(\mathcal{G}) \leq \tilde{\omega}(\mathcal{G}),$$

with equality holding for trees. The following question arises.

QUESTION **1** *On which graphs $\mathcal{G}$ (if any) do we obtain strict inequality $\omega(\mathcal{G}) < \tilde{\omega}(\mathcal{G})$?*

Aside from Γ_n and $\tilde{\Gamma}_n$, a variety of other subsets of Σ_n enjoy (3.2) and (3.3). Whenever either some specific geometric object or pattern does not show up along a path of length n, for example, a self-intersection at a vertex of $\mathcal{G}$, a self-intersection at an edge of $\mathcal{G}$, a certain *rectangle,* a certain *triple-loop,* or a *self-intersection with a certain geometric object* constructed from the SRW-path, then concatenating two paths may or may not preserve the defining property. As a result, an inequality of type (3.1) emerges. This also means that, by suppressing such geometric events, other *generalized weakly self-avoiding walks* may be defined in place of the ones with measures $\mathbf{Q}_n^\beta$ and $\tilde{\mathbf{Q}}_n^\beta$, respectively.

The features of J_n and $\tilde{J}_n$ reported here may bear a crucial impact on the values of the critical exponent "displacement exponent" of the weakly SAW and weakly SAT on graphs with small degree (this is suggested by the analysis and results in HUETER [4]). Therefore, if we think of the *displacement exponent* or the *root mean square displacement exponent* $\nu_\beta(\mathcal{G})$ and $\tilde{\nu}_\beta(\mathcal{G})$ of the weakly SAW and weakly SAT, respectively, being defined by

$$\begin{aligned} \nu_\beta(\mathcal{G}) &= \lim_{n\to\infty} \ln \mathbf{E}_\beta \|S_n\|^2/(2\ln n), \\ \tilde{\nu}_\beta(\mathcal{G}) &= \lim_{n\to\infty} \ln \tilde{\mathbf{E}}_\beta \|S_n\|^2/(2\ln n), \end{aligned}$$

respectively, whenever the limits exist (otherwise, we consider the corresponding upper and lower exponents), we are led to

Question **2** *Which classes of generalized weakly self-avoiding walks in* $\mathbf{Z}^2$ *exhibit a displacement exponent* $\nu^*_\beta(\mathbf{Z}^2) = 3/4$ *?*

Question **3** *What classes of generalized weakly self-avoiding walks on* $\mathcal{G}$ *have displacement exponent* $\nu^*_\beta(\mathcal{G}) = \nu_\beta(\mathcal{G})$ *or* $\nu^*_\beta(\mathcal{G}) = \tilde{\nu}_\beta(\mathcal{G})$*, the same exponent as the weakly SAW or weakly SAT on* $\mathcal{G}$ *?*

4 Number of Self-Intersection Events

The models that put a price on each self-intersection event focus their attention on the paths that exhibit a significantly smaller number of self-intersection events than expected of the SRW for small degree of the graph. If a self-intersection event occurs, it occurs often but not too often. We shall begin with a few elementary inequalities that lie at the core of later estimates. We stress that the results which we derive here apply to both measures $\mathbf{Q}_n^\beta$ and $\tilde{\mathbf{Q}}_n^\beta$. We carry out the proofs for the weakly SAW and do not repeat it for the weakly SAT here. The adjustments needed in the arguments when switching from one model to the other are minor.

Let A_1, A_2 denote any two subsets of $[0, n^2]$. From the definition of the measure $\mathbf{Q}_n^\beta$ in (2.3) in connection with $\mathbf{E}_0 \exp\{-\beta J_n\} \geq \mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_2\}})$, we derive

$$\begin{aligned} \mathbf{Q}_n^\beta(J_n \in A_1) &= \frac{\mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_1\}})}{\mathbf{E}_0(\exp\{-\beta J_n\})} \\ &\leq \frac{\mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_1\}})}{\mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_2\}})}. \end{aligned} \tag{4.1}$$

As a consequence of (4.1), if for any integer $n \geq 1$ and any $A_1 \subset [0, n^2]$, we come up with a subset A_2 of $[0, n^2]$ and some real number $\tau_* > 0$ that does not depend on n such that

$$\mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_1\}}) \;<\; e^{-\tau_* n} \cdot \mathbf{E}_0(\exp\{-\beta J_n\}\, 1_{\{J_n \in A_2\}}), \tag{4.2}$$

then it follows from (4.1) that

$$\mathbf{Q}_n^\beta(J_n \in A_1) < e^{-\tau_* n}. \tag{4.3}$$

This tells us that, when for some set A_2 there is a subset of Σ_n with $J_n \in A_2$ which contributes strictly more to $\mathbf{E}_0 \exp\{-\beta J_n\}$ than the subset of Σ_n with $J_n \in A_1$ in the sense of inequality (4.2), then the event $\{J_n \in A_1\}$ has a $\mathbf{Q}_n^\beta$-probability that is exponentially decaying in n. Hence, for large n, with overwhelming probability, $J_n \notin A_1$. We will fix A_1 and find a suitable set A_2 along with a τ_* in order to determine likely upper and lower bounds on J_n.

Proposition **1** (Upper Bound on J_n) *Assume that* $\mathcal{G}$ *is an infinite, locally finite connected transitive graph of degree* $deg(\mathcal{G}) \geq 2$*. Let* $\omega(\mathcal{G})$ *denote the exponent of the number of self-avoiding walks, as stated in* (3.2)*. Let* $\beta > 0$ *and choose any suitably small* $\epsilon > 0$*. If*

$\tau_u = \tau_u(\mathcal{G}) = \epsilon\,(\ln(deg(\mathcal{G}))-\omega(\mathcal{G})) > 0$ *and* $b_2 = b_2(\mathcal{G},\beta) = (1+\epsilon)\,(\ln(deg(\mathcal{G}))-\omega(\mathcal{G}))/\beta > 0$, *then for every* $b \geq b_2$ *and every integer* $n \geq 1$,

$$\mathbf{Q}_n^\beta(J_n > bn) < e^{-\tau_u\, n}.$$

Proof. We write $d = \deg(\mathcal{G})$ and $\omega = \omega(\mathcal{G})$. Let $\beta > 0$, fix any suitably small $\epsilon > 0$, and choose $b \geq b_2(\mathcal{G},\beta) = (1+\epsilon)\,(\ln d - \omega)/\beta > 0$. Recalling that $\mathbf{P}_0$ denotes probability relative to the SRW, we collect

$$\mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n > bn\}}) \;<\; e^{-\beta bn}\, \mathbf{P}_0(J_n > bn) \leq e^{-\beta bn}. \tag{4.4}$$

In addition, in view of (3.3),

$$\mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n = 0\}}) \;=\; \mathbf{P}_0(J_n = 0) \geq e^{n\omega}\, d^{-n} = e^{-n(\ln d - \omega)}. \tag{4.5}$$

Combining (4.4) and (4.5) with $-\beta b \leq -(1+\epsilon)\,(\ln d - \omega) < -(\ln d - \omega)$, we arrive at

$$\begin{aligned}
\mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n > bn\}}) &< e^{-\beta bn} \\
&\leq e^{-n\,(1+\epsilon)\,(\ln d - \omega)} \\
&= e^{-n\,\epsilon\,(\ln d - \omega)} \cdot e^{-n\,(\ln d - \omega)} \\
&\leq e^{-\tau_u\, n} \cdot \mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n = 0\}})
\end{aligned} \tag{4.6}$$

for $\tau_u = \tau_u(\mathcal{G}) = \epsilon\,(\ln d - \omega) > 0$, which does not depend on β. Now, the advertized claim follows from (4.2) with $A_1 = (bn, n^2]$, $A_2 = \{0\}$, and $\tau_* = \tau_u(\mathcal{G})$. □

An instance of the result is when $\mathcal{G} = \mathbf{Z}^d$ for $d \geq 1$, and thus, $\deg(\mathcal{G}) = 2d \geq 2$. Parallel reasoning to the one employed in the last proof yields that, if $\tilde{\tau}_u = \tilde{\tau}_u(\mathcal{G}) = \epsilon\,(\ln(\deg(\mathcal{G})) - \tilde{\omega}(\mathcal{G})) > 0$ (with $\tilde{\omega}(\mathcal{G})$ defined in (3.7)) and $\tilde{b}_2 = \tilde{b}_2(\mathcal{G},\beta) = (1+\epsilon)\,(\ln(\deg(\mathcal{G}))-\tilde{\omega}(\mathcal{G}))/\beta > 0$, then for every $b \geq \tilde{b}_2$ and every integer $n \geq 1$,

$$\tilde{\mathbf{Q}}_n^\beta(\tilde{J}_n > bn) < e^{-\tilde{\tau}_u\, n}. \tag{4.7}$$

We turn to estimate the $\mathbf{Q}_n^\beta$-probability of the paths in Σ_n with small J_n and to prove a likely lower bound on J_n of the same order of magnitude in n. This is quite a bit more subtle than the upper bound.

PROPOSITION **2** (Lower Bound on J_n) *Assume that* $\mathcal{G}$ *is an infinite, locally finite connected transitive graph of degree* ≥ 2. *Let* $\beta > 0$. *There are positive finite constants* $\hat{\alpha}, b_*(\beta) > 2\hat{\alpha}$, *and* $\tau_l = \tau_l(\mathcal{G},\beta)$ *(all specified below), independent of* n *but dependent on* $\mathcal{G}$ *and* β, *such that for every* $0 < b_1 < b_*(\beta) - 2\hat{\alpha}$ *and every sufficiently large integer* n,

$$\mathbf{Q}_n^\beta(J_n < b_1 n) < e^{-\tau_l\, n}.$$

Proof. As before, our goal is to apply (4.2) and (4.3), this time with $A_1 = [0, b_1 n)$ and $A_2 = [0, bn]$ for some suitable $b > b_1$ and appropriate τ_*. To warm up on our calculation ahead, imagine the following. If we pick any path in Γ_n (with $J_n = 0$), distribute bn self-intersections among its sites by introducing "backtracking steps," and count these newly generated paths, then an elementary counting exercise shows that, at least for suitably small b, the set of these created paths contributes significantly more to $\mathbf{E}_0(\exp\{-\beta J_n\})$ than Γ_n (since the number of paths created overcompensates the penalizing weight). This comparison indeed works in the same fashion if we start with paths that exhibit some self-intersections but not too many, even though to carry out the proof, we need to be careful with overlap of the added self-intersections with self-intersections that exist prior to the alteration of paths and with creating paths multiple times.

Write $d = \deg(\mathcal{G})$ and let $\beta > 0$ be fixed. Pick a suitably large finite number $M > 2$ and a suitably small $0 < \hat{\alpha} < (2d)^{-2M}$. Recall $b_2(\mathcal{G}, \beta)$ from Proposition 1. We fix the following notation (most of it will not be used until the end of the proof)

$$\begin{aligned} \kappa &= \hat{\alpha} M \ln(2d) \qquad (4.8)\\ \zeta_*(\beta) &= \zeta_*(\mathcal{G},\beta) = \min\{\kappa,\ \beta M\hat{\alpha},\ \beta b_2(\mathcal{G},\beta),\ 4\beta(1/2 - e^{-1})/(6+d)\}\\ b_*(\beta) &= b_*(\mathcal{G},\beta) = \zeta_*(\beta)/\beta. \end{aligned}$$

We will assume that $2\hat{\alpha} < b \le b_*(\beta)$ and $0 < b_1 < b - 2\hat{\alpha}$. In particular, we see that $b \le 4(1/2 - e^{-1})/(6+d)$.

Recall the set Σ_n of SRW-paths $\omega = S_0, S_1, \ldots, S_n$ on $\mathcal{G}$ of length n, starting at $\mathbf{0}$, with their SILT $J_n = J_n(S_0, S_1, \ldots, S_n)$ (We will simply write J_n). For any $0 < s < b$ and any integer $n \ge 1$, define the sets

$$\begin{aligned} \Omega_n^{b_1} &= \{\omega \in \Sigma_n : J_n < b_1 n\} \qquad (4.9)\\ \Lambda_n^s &= \{\omega \in \Sigma_n : J_n \in \{\lfloor sn \rfloor - 2, \lfloor sn \rfloor - 1, \lfloor sn \rfloor\}\}, \end{aligned}$$

where $\lfloor sn \rfloor$ denotes the integer part of sn. It is evident that

$$\begin{aligned} \mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n < b_1 n\}}) &\le \mathbf{P}_0(J_n < b_1 n)\\ &= \exp\{-n\,(\ln d - \ln|\Omega_n^{b_1}|/n)\}. \qquad (4.10) \end{aligned}$$

Finding a *lower* bound on $\mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n < bn\}})$ for all $b > b_1$ and determining when this lower bound *strictly* exceeds the righthand side of (4.10) suffices to accomplish our proof. As we have seen earlier, once b is larger than a certain value, the penalizing weight $e^{-\beta J_n}$ becomes too small for the paths in Σ_n with $J_n > bn$ to significantly contribute to $\mathbf{E}_0 e^{-\beta J_n}$. Since $bn \ge \lfloor sn \rfloor$ for $0 < s < b$, by inspection,

$$\begin{aligned} \mathbf{E}_0(e^{-\beta J_n}\, 1_{\{J_n < bn\}}) &> e^{-\beta bn}\, \mathbf{P}_0(J_n < bn)\\ &\ge e^{-\beta bn}\, \mathbf{P}_0(J_n \in \{\lfloor sn \rfloor - 2, \lfloor sn \rfloor - 1, \lfloor sn \rfloor\})\\ &= e^{-\beta bn}\, |\Lambda_n^s|\, d^{-n}. \qquad (4.11) \end{aligned}$$

The bulk of the proof consists in finding a lower bound on $|\Lambda_n^s|$. To this end, we pick certain paths γ in Σ_m for a suitable step size $m = m(\gamma) < n$ which exhibit $J_m \le b_1 n$ and – with the help of a scheme to place repetition steps in γ and add self-intersections to the walk – construct paths in Λ_n^s for $s < b$. Estimating from below the number of constructed paths will thus lead to desired lower bounds on $|\Lambda_n^s|$ and $\mathbf{E}_0(e^{-\beta J_n} 1_{\{J_n < bn\}})$.

Repetition steps. We turn to describe two kinds of "repetition steps" to be implemented along a path $\gamma \in \Sigma_m$ and a scheme to place those repetition steps in order to produce a path $\tilde{\gamma} \in \Lambda_n^s$. We say that a path $\gamma \in \Sigma_m$ has a *single repetition step* (SR) at x_* if

$$\begin{aligned} S_{j+1} &= S_{j-1} && \text{(immediate backtracking from } x_*) \\ S_{j+2} &= S_j = x_* && \text{(moving on to } x_*) \end{aligned} \tag{4.12}$$

for some integer $1 \le j \le m-2$ (half a single repetition step will mean $S_j = x_*$ and $S_{j+1} = S_{j-1}$) and that $\gamma \in \Sigma_m$ has a *double repetition step* (DR) at x_{**} if

$$\begin{aligned} S_{i+3} = S_{i+1} &= S_{i-1} && \text{(backtracking from } x_{**}) \\ S_{i+4} = S_{i+2} &= S_i = x_{**} && \text{(moving on to } x_{**}) \end{aligned} \tag{4.13}$$

for some integer $1 \le i \le m-4$. While the locations (on γ) of the double repetition steps are *prescribed* once the particular γ has been chosen, the locations of the single repetition steps are not in that there is a variety of possible ways to select them. Specifically, we prescribe a double repetition step at each site x_* of γ for which there is an integer $1 \le k$ such that

$$\begin{aligned} &S_{k+1} = S_{k-1} \quad \text{and} \quad S_{k+2} = S_k = x_* \text{ and} && \text{(SR at } x_*) \\ &\nexists\, 0 < j \notin \{k, k+2\} \text{ with } S_j = S_k \text{ or} && \text{(no other visits to } x_* \text{ or)} \\ &\nexists\, 0 < j' \notin \{k-1, k+1\} \text{ with } S_{j'} = S_{k-1} && \text{(no other visits to } S_{k-1}). \end{aligned} \tag{4.14}$$

If $d(\gamma)$ denotes the number of self-intersections that the placed DRs add to those of γ, then we choose γ and a positive real number $\alpha(\gamma) \ge \hat{\alpha} + 1/n$ that depends on the path γ satisfying $2\alpha(\gamma)n \in \mathbf{Z}$ such that, for $s \ge b_1 + 2\hat{\alpha}$,

$$\begin{aligned} J_{n(1-2\alpha(\gamma))} + d(\gamma) &< b_1 n \\ J_{n(1-2\alpha(\gamma))} + d(\gamma) + 2\,\alpha(\gamma)\,n &\in \{\lfloor sn \rfloor - 1, \lfloor sn \rfloor\}. \end{aligned} \tag{4.15}$$

We thus take $\gamma \in \Sigma_{n(1-2\alpha(\gamma))}$ for some suitable $\alpha(\gamma)$. Write $\mathcal{K}_n(b_1)$ for the set of paths γ that possess the properties stated in (4.15). In particular, (4.15) implies that, for any $\gamma \in \mathcal{K}_n(b_1)$, its SILT $J_{n(1-2\alpha(\gamma))}(\gamma)$ satisfies

$$J_{n(1-2\alpha(\gamma))}(\gamma) < b_1 n - d(\gamma) \le b_1 n. \tag{4.16}$$

Note that $\mathcal{K}_n(b_1)$ depends on the collection $\{\alpha(\gamma)\}$ and in that sense depends on $\hat{\alpha}$ (yet we suppress this dependency when writing $\mathcal{K}_n(b_1)$). If we write $s_*(n)n$ for the number arising on the righthand side of the second line in (4.15), that is, $s_*(n) = (\lfloor sn \rfloor - 1)/n$ or $s_*(n) = \lfloor sn \rfloor/n \le s < b$ so that $s_*(n)n = \lfloor sn \rfloor - 1$ or $s_*(n)n = \lfloor sn \rfloor$, respectively, and solve the second line in (4.15) to get

$$\alpha(\gamma) = (s_*(n)n - J_{n(1-2\alpha(\gamma))} - d(\gamma))/(2n) \tag{4.17}$$

with $\alpha(\gamma) \le s/2 < b/2$ and $\alpha(\gamma)n \in \mathbf{Z}/2$, then with

$$\alpha_*(\gamma) = \lfloor \alpha(\gamma)n \rfloor / n$$

we see that $\alpha_*(\gamma)n \in \mathbf{Z}$ and $\alpha_*(\gamma) \ge \hat{\alpha}$.

Scheme to place single repetition steps. For any $b_1 + 2\hat{\alpha} \le s < b$ and $\gamma \in \mathcal{K}_n(b_1)$ together with a suitable $\alpha(\gamma)$, we now choose $\alpha_*(\gamma) \cdot n$ suitable locations on the path γ to place $\alpha_*(\gamma) \cdot n$ SR in order to construct a new path $\tilde{\gamma}$ in Λ_n^s. To obtain the path $\tilde{\gamma}$ from γ, we think of walking along the path γ, starting at $\mathbf{0}$, and at each location, marked for a SR or DR, of carrying out that repetition step. For any path $\gamma \in \Sigma_m$, denote the set

$$\mathcal{R}_\gamma = \{x \in \gamma : x = S_i = S_j \text{ for some } 0 \le i < j \le m\}$$

of sites that the walk along γ visited more than once. If $R_\gamma(x_*) = \{S_{j-1}, x_*\}$ denotes the two sites, highlighted in (4.12), that are revisited during the SR at x_*, define

$$\mathcal{A}_\gamma = \{x \in \gamma : R_\gamma(x) \cap \mathcal{R}_\gamma = \emptyset\}.$$

For any $b_1 + 2\hat{\alpha} \le s < b$ and $\gamma \in \mathcal{K}_n(b_1)$ with a suitable $\alpha(\gamma)$, we choose $\alpha_*(\gamma) \cdot n$ *distinct* sites $x_1, x_2, \ldots, x_{\alpha_*(\gamma)\cdot n}$ among the sites in $\mathcal{A}_\gamma$ such that the $R_\gamma(x_k)$ are *pairwise disjoint* sets, thus,

$$\begin{array}{rcll} x_k & \in & \mathcal{A}_\gamma & \forall\, 1 \le k \le \alpha_*(\gamma) \cdot n \quad \text{(locations of placed SRs)}, \\ R_\gamma(x_k) & \cap & R_\gamma(x_l) = \emptyset & \text{for } k \ne l \quad \text{(separation of SRs)}. \end{array} \tag{4.18}$$

For any $\gamma \in \mathcal{K}_n(b_1)$ combined with a vector of sites $x_1, x_2, \ldots, x_{\alpha_*(\gamma)\cdot n}$ enjoying (4.18), we end up with a new path $\tilde{\gamma} \in \Lambda_n^s$. The path $\tilde{\gamma}$ is uniquely determined by s, γ along with its prescribed DRs, and the scheme to place the SRs at $x_1, x_2, \ldots, x_{\alpha_*(\gamma)\cdot n}$. If we write $\Phi^\circ(\cdot;\cdot)$ for the map which maps $(\gamma; x_1, \ldots, x_{\alpha_*(\gamma)\cdot n})$ onto $\tilde{\gamma}$ and thus satisfies

$$\Phi^\circ(\gamma; x_1, \ldots, x_{\alpha_*(\gamma)\cdot n}) = \tilde{\gamma},$$

then for any $\gamma \in \mathcal{K}_n(b_1)$, define

$$G_\gamma = \{\tilde{\gamma} \in \Lambda_n^s : \tilde{\gamma} = \Phi^\circ(\gamma; x_1, \ldots, x_{\alpha_*(\gamma)\cdot n}) \text{ for some } x_1, \ldots, x_{\alpha_*(\gamma)\cdot n} \text{ obeying (4.18)}\}.$$

Rule (4.18) makes sure that placed SRs do not overlap with one another and that each SR placed on γ not overlap with any self-intersection of γ or with any of the self-intersections introduced by the DRs because the DRs only add self-intersections at sites $R_\gamma(x_*)$ for some x_* that carries a SR. Hence, rule (4.18) guarantees that each of the SR described in

(4.18) create two self-intersections of $\tilde{\gamma}$ which γ does not have, with one possible exception when there is half a SR, adding one self-intersection. As a consequence, the number of self-intersections of $\tilde{\gamma}$ equals

$$J_{n(1-2\alpha(\gamma))} + d(\gamma) + 2\alpha(\gamma)n \quad \text{or} \quad J_{n(1-2\alpha(\gamma))} + d(\gamma) + 2\alpha(\gamma)n - 1,$$

which, in view of (4.15), is a number in $\{\lfloor sn\rfloor - 2, \lfloor sn\rfloor - 1, \lfloor sn\rfloor\}$. Hence, $\tilde{\gamma} \in \Lambda_n^s$.

Moreover, it is apparent from (4.18) that, aside from $x_k \notin \mathcal{R}_\gamma$ for all $1 \le k \le \alpha_*(\gamma)\cdot n$, also for all $1 \le k \le \alpha_*(\gamma)\cdot n$, we have $x_k \notin \{S_{i+1}, S_{j+1}\}$ for any $S_i = S_j \in \mathcal{R}_\gamma$. Since the walk may use two different edges to arrive at and leave some particular site of γ, we conclude that excluding half (or $\lfloor d/2\rfloor$) of the neighboring sites of any site in $\mathcal{R}_\gamma$, gives a set of sites which is contained in the set $\mathcal{A}_\gamma$ of sites that are possible choices to place the SRs. Keeping in mind that the length or the number of edges of $\gamma \in \mathcal{K}_n(b_1)$ equals $n(1-2\alpha(\gamma))$ and realizing that $|\mathcal{R}_\gamma| < b_1 n$ by virtue of (4.16) (some sites of γ that are not free of self-intersections may have more than one self-intersection), we are led to the lower bounds on the size of $\mathcal{A}_\gamma$,

$$\begin{aligned} |\mathcal{A}_\gamma| &\ge n(1-2\alpha(\gamma)) - |\mathcal{R}_\gamma|\, d/2 \\ &> n(1-2\alpha(\gamma) - db_1/2) \ge \lfloor n(1-2\alpha(\gamma) - db_1/2)\rfloor \end{aligned}$$

and

$$\lfloor |\mathcal{A}_\gamma|/2\rfloor \ \ge\ \lfloor n(1/2-\alpha(\gamma)-db_1/4)\rfloor. \tag{4.19}$$

Counting. In order to find a lower bound on the count of paths in G_γ, two moments' thoughts reveal that the set of all ways to select the sites $x_1, x_2, \dots, x_{\alpha_*(\gamma)\cdot n}$ stands in one-to-one correspondence with the set G_γ. Its cardinality $|G_\gamma|$ equals the number of ways to distribute $\alpha_*(\gamma)\cdot n$ identical balls into $|\mathcal{A}_\gamma|$ urns under the restrictions that "no urn contains more than one ball" and "at most one of two urns associated with two consecutive sites of γ contains a ball." In view of the second restriction, the $\alpha_*(\gamma)\cdot n$ balls are distributed among at least $\lfloor |\mathcal{A}_\gamma|/2\rfloor$ urns. Hence, in light of (4.19), setting $a(\gamma) = \lfloor n(1/2-\alpha(\gamma)-db_1/4)\rfloor/n$ (thus, $a(\gamma)n \in \mathbf{Z}$) and appealing to the Stirling approximation $k! = \sqrt{2\pi k}\, e^{-k}k^k(1+o(1))$ as $k\to\infty$ (the notation $f(x) = o(g(x))$ as $x\to\infty$ for two real-valued functions f and g means that $\lim_{x\to\infty} f(x)/g(x) = 0$), we collect

$$\begin{aligned} |G_\gamma| &\ge \binom{\lfloor |\mathcal{A}_\gamma|/2\rfloor}{\alpha_*(\gamma)\cdot n} \\ &\ge \binom{\lfloor n(1/2-\alpha(\gamma)-db_1/4)\rfloor}{\alpha_*(\gamma)\cdot n} \\ &= \binom{a(\gamma)\cdot n}{\alpha_*(\gamma)\cdot n} \\ &= (1+o(1))\left(\frac{a(\gamma)^{a(\gamma)}}{\alpha_*(\gamma)^{\alpha_*(\gamma)}\,(a(\gamma)-\alpha_*(\gamma))^{a(\gamma)-\alpha_*(\gamma)}}\right)^n \left(\frac{1}{2\pi n}\right)^{1/2} \\ &\quad \cdot \left(\frac{a(\gamma)}{\alpha_*(\gamma)\,(a(\gamma)-\alpha_*(\gamma))}\right)^{1/2} \end{aligned} \tag{4.20}$$

as $n \to \infty$. We claim that

$$\frac{a(\gamma)^{a(\gamma)}}{(a(\gamma)-\alpha_*(\gamma))^{a(\gamma)-\alpha_*(\gamma)}} > 1 \tag{4.21}$$

for all sufficiently large n. To see this, consider the functions $f(x) = x \ln x$ and $g(x) = x^x = \exp\{f(x)\}$ for $x \in (0,1]$. Then $f(b) < f(c)$ for some $b, c \in (0,1]$ is equivalent to $g(b) < g(c)$. The function f has derivative $f'(x) = \ln x + 1$ and f is strictly decreasing on $(0, e^{-1})$ and strictly increasing on $(e^{-1}, 1]$. Furthermore, f has a local minimum at $x = e^{-1}$, $f(e^{-1}) = -e^{-1}$ and $\lim_{x\to 0+} f(x) = 0 = f(1)$. Hence, if $e^{-1} \le b < c$, we have $f(b) < f(c)$, equivalently, $g(b) < g(c)$. Therefore, as long as $a(\gamma) - \alpha_*(\gamma) \ge e^{-1}$, we set $b = a(\gamma) - \alpha_*(\gamma)$ and $c = a(\gamma)$ to obtain $g(c)/g(b) > 1$, and thus, (4.21). But when $1/n \le \alpha(\gamma)$, then (recall $\alpha_*(\gamma) \le \alpha(\gamma) < b/2$ and $b_1 < b$),

$$\begin{aligned} a(\gamma) - \alpha_*(\gamma) &\ge \lfloor n(1/2 - 2\alpha(\gamma) - db_1/4) \rfloor / n \\ &\ge 1/2 - 2\alpha(\gamma) - db_1/4 - 1/n \\ &\ge 1/2 - 3\alpha(\gamma) - db_1/4 \\ &> 1/2 - 3b/2 - db/4 = 1/2 - b(6+d)/4, \end{aligned}$$

which is $\ge e^{-1}$ since we assumed that $b \le 4(1/2 - e^{-1})/(6+d)$. We conclude that for all sufficiently large n, we have $a(\gamma) - \alpha_*(\gamma) \ge e^{-1}$ and thus verified (4.21).

In addition, because the function $h(x) = 1/g(x) = 1/x^x$ is increasing on $(0, e^{-1})$ and our assumption $\alpha_*(\gamma) \ge \hat\alpha > 0$, it follows that $1/\alpha_*(\gamma)^{\alpha_*(\gamma)} \ge 1/\hat\alpha^{\hat\alpha} > 1$. This together with (4.20) and (4.21) implies that there is some $\xi_{\hat\alpha} > 0$, depending on $\hat\alpha$ but not on γ, with

$$\exp\{\xi_{\hat\alpha}\} > 1/\hat\alpha^{\hat\alpha} > 1$$

such that for all sufficiently large n,

$$|G_\gamma| \quad \ge \quad \exp\{\xi_{\hat\alpha}\, n\}. \tag{4.22}$$

To estimate the number of paths in $G_\gamma \subset \Lambda_n^s$ that do not lie in $\cup_{\gamma' \ne \gamma \in \mathcal{K}_n(b_1)} G_{\gamma'}$, consider the intersection $G_\gamma \cap G_{\gamma'}$ for two distinct paths $\gamma, \gamma' \in \mathcal{K}_n(b_1)$, that is, $\gamma \ne \gamma'$. If for any two distinct paths $\gamma, \gamma' \in \mathcal{K}_n(b_1)$, one of the paths γ and γ' has a SR at x_* such that a DR is prescribed and the other of the two paths has an existing SR and DR at x_*, a "triple repetition step" (so that, of course, no DR is prescribed), and $\alpha_*(\gamma)$ and $x_1, x_2, \ldots, x_{\alpha_*(\gamma)\cdot n}$ are identical for both γ and γ', then possibly $\tilde\gamma = \tilde\gamma'$. However, the event $\tilde\gamma = \tilde\gamma'$ does not arise otherwise.

For any $\gamma \in \mathcal{K}_n(b_1)$, let $\mathcal{F}(\gamma) = \{x_* \in \gamma : x_* \text{ has a SR or } x_* \text{ has a SR and a DR}\}$. If a path $\tilde\gamma$ is in $G_\gamma \cap G_{\gamma'}$ for $\gamma \ne \gamma'$, then

(i) $\alpha_*(\gamma)$ and $x_1, x_2, \ldots, x_{\alpha_*(\gamma)\cdot n}$ are identical for both γ and γ',

(ii) γ and γ' can only differ on $\mathcal{F}(\gamma)$, and

(iii) γ and γ' must differ at least at one site of $\mathcal{F}(\gamma)$.

For fixed γ, each path γ' that together with γ has the features (i)–(iii) and satisfies

$\Phi^\circ(\gamma; x_1, \ldots, x_{\alpha_*(\gamma)\cdot n}) = \Phi^\circ(\gamma'; x_1, \ldots, x_{\alpha_*(\gamma)\cdot n})$ for at least one $x_1, \ldots, x_{\alpha_*(\gamma)\cdot n}$ generates a set $G_{\gamma'} = G_\gamma$ (thus, all images under Φ° of γ and γ' coincide). In particular, for any two γ and γ' in $\mathcal{K}_n(b_1)$, we see that either

$$G_{\gamma'} = G_\gamma \quad \text{or} \quad G_{\gamma'} \cap G_\gamma = \emptyset.$$

For $\gamma \in \mathcal{K}_n(b_1)$, we count the paths in $\mathcal{H}(\gamma) = \{\gamma' \in \mathcal{K}_n(b_1) : G_{\gamma'} = G_\gamma\}$ as follows: at all but one site $x_* \in \mathcal{F}(\gamma)$ at which γ and γ' need to differ, there are two possibilities of γ', either to have a SR or a SR and a DR. Thus, since each SR and each SR and DR both produce at least 2 self-intersections, one minute's thoughts reveal that $|\mathcal{F}(\gamma)| < b_1 n/2$ thanks to (4.16). We conclude that for any $\gamma \in \mathcal{K}_n(b_1)$,

$$|\mathcal{H}(\gamma)| \leq 2^{\,b_1 n/2}\,.$$

Note that $\cup_{\gamma\in\mathcal{K}_n(b_1)} G_\gamma \subset \Lambda_n^s$ and that the inclusion is strict (there are paths in Λ_n^s which are not in $\cup_{\gamma\in\mathcal{K}_n(b_1)} G_\gamma$ because our repetition steps do not include loops). Pulling these steps together, we find that for $s \geq b_1 + 2\hat{\alpha}$,

$$\sum_{\gamma\in\mathcal{K}_n(b_1)} |G_\gamma| \leq |\Lambda_n^s| \cdot 2^{b_1 n/2}\,.$$

In light of (4.22),

$$\sum_{\gamma\in\mathcal{K}_n(b_1)} |G_\gamma| \geq \exp\{\xi_{\hat{\alpha}}\, n\} \cdot |\mathcal{K}_n(b_1)|$$

for all sufficiently large n. Merging these inequalities yields for all sufficiently large n,

$$|\Lambda_n^s| \quad \geq \quad \exp\{\xi_{\hat{\alpha}}\, n\} \cdot |\mathcal{K}_n(b_1)| \cdot 2^{-b_1 n/2}\,. \tag{4.23}$$

Estimation of $\tau_l(\mathcal{G}, \beta)$. Fix any $0 < \epsilon < (s/2)(\ln 2/\ln d) < b/2$. We set

$$\tau_l(\mathcal{G}, \beta) = \epsilon\,\beta b > 0. \tag{4.24}$$

Thus, $\tau_l(\mathcal{G}, \beta)$ does not depend on n and $\epsilon\,\beta b_1 < \tau_l(\mathcal{G}, \beta) \leq (b/2)\beta b = \beta b^2/2$. Note that

$$\beta b + \ln d - \ln|\Lambda_n^s|/n \geq \beta b > 0 \tag{4.25}$$

because $\ln d \geq \ln|\Lambda_n^s|/n$. If we write $F_n = \beta b + \ln d - \ln|\mathcal{K}_n(b_1)|/n - \xi_{\hat{\alpha}} + (\ln 2)b_1/2$ and verify the claim that

$$(1+\epsilon)F_n < \ln d - \ln|\Omega_n^{b_1}|/n, \tag{4.26}$$

then, by virtue of (4.10), (4.11), (4.23), (4.24), and (4.25) we arrive at

$$\begin{aligned} \mathbf{E}_0(e^{-\beta J_n}\,1_{\{J_n<bn\}}) &\overset{(4.11)}{>} \exp\{-n\,(\beta b - \ln|\Lambda_n^s|/n + \ln d)\} \\ &= \exp\{-n\,(\beta b - \ln|\Lambda_n^s|/n + \ln d)(1+\epsilon-\epsilon)\} \\ &= \exp\{n\epsilon\,(\beta b - \ln|\Lambda_n^s|/n + \ln d)\} \end{aligned} \tag{4.27}$$

$$
\begin{aligned}
&\quad \cdot \exp\{-n\,(\beta b - \ln|\Lambda_n^s|/n + \ln d)(1+\epsilon)\}\\
&\overset{(4.25)}{\geq} \exp\{n\epsilon\beta b\}\cdot\exp\{-n\,(\beta b - \ln|\Lambda_n^s|/n + \ln d)(1+\epsilon)\}\\
&\overset{(4.24)}{=} \exp\{n\,\tau_l(\mathcal{G},\beta)\}\cdot\exp\{-n\,(\beta b + \ln d - \ln|\Lambda_n^s|/n)(1+\epsilon)\}\\
&\overset{(4.23)}{\geq} \exp\{n\,\tau_l(\mathcal{G},\beta)\}\\
&\quad \cdot \exp\{-n\,(\beta b + \ln d - \ln|\mathcal{K}_n(b_1)|/n - \xi_{\hat\alpha} + (\ln 2)b_1/2)(1+\epsilon)\}\\
&= \exp\{n\,\tau_l(\mathcal{G},\beta)\}\cdot\exp\{-n\,F_n\,(1+\epsilon)\}\\
&\overset{(4.26)}{>} \exp\{n\,\tau_l(\mathcal{G},\beta)\}\cdot\exp\{-n\,(\ln d - \ln|\Omega_n^{b_1}|/n)\}\\
&\overset{(4.10)}{\geq} \exp\{n\,\tau_l(\mathcal{G},\beta)\}\cdot\mathbf{E}_0(e^{-\beta J_n}\,1_{\{J_n<b_1 n\}})
\end{aligned}
$$

for $b_1 + 2\hat\alpha < b \leq b_2(\mathcal{G},\beta)$ (for $b_2(\mathcal{G},\beta)$ as in Proposition 1) and all sufficiently large n, equivalently,

$$
\begin{aligned}
\mathbf{E}_0(e^{-\beta J_n}\,1_{\{J_n<b_1 n\}}) &\overset{(4.27)}{<} \exp\{-n\,\tau_l(\mathcal{G},\beta)\}\cdot\mathbf{E}_0(e^{-\beta J_n}\,1_{\{J_n<bn\}})\\
&\leq \exp\{-n\,\tau_l(\mathcal{G},\beta)\}\cdot\mathbf{E}_0(e^{-\beta J_n}\,1_{\{J_n\leq bn\}}). \qquad (4.28)
\end{aligned}
$$

As a result, once we show (4.26), letting $A_1 = [0, b_1 n)$, $A_2 = [0, bn]$, and $\tau_* = \tau_l(\mathcal{G},\beta)$ and applying (4.2) and (4.3) will achieve our proof.

Proof of (4.26). Inequality (4.26) is tantamount to

$$\ln|\mathcal{K}_n(b_1)|/n - (1+\epsilon)^{-1}\ln|\Omega_n^{b_1}|/n + \xi_{\hat\alpha} - (\ln 2)b_1/2 - \epsilon(1+\epsilon)^{-1}\ln d > \beta b. \qquad (4.29)$$

We will show that the lefthand side of (4.29) is strictly bounded below by some real number $\kappa > 0$. We thus can let $\beta b \leq \kappa$. First, note that any path $\gamma \in \mathcal{K}_n(b_1)$ may be completed to give a path in $\Omega_n^{b_1}$. There are no more than $d^{2\alpha(\gamma)n}$ ways to complete γ. Recalling that $2\alpha(\gamma) < s$ (see (4.17)), we get

$$|\Omega_n^{b_1}| \leq |\mathcal{K}_n(b_1)|\cdot d^{sn},$$

equivalently, $(\ln|\Omega_n^{b_1}| - \ln|\mathcal{K}_n(b_1)|)/n \leq s\ln d$. This implies that

$$\frac{1}{n}(\ln|\mathcal{K}_n(b_1)| - (1+\epsilon)^{-1}\ln|\Omega_n^{b_1}| \geq -s\ln d.$$

But $-(\ln 2)b_1/2 > -(\ln 2)s/2$ because $b_1 < s$ and $-\epsilon(1+\epsilon)^{-1}\ln d > -s\ln 2/2$ because $\epsilon < (s/2)(\ln 2/\ln d)$. In addition, we assumed in (4.8) that $s < b \leq \hat\alpha M$, thus, $-s > -\hat\alpha M$. Assembling these estimates,

$$
\begin{aligned}
\ln|\mathcal{K}_n(b_1)|/n - (1+\epsilon)^{-1}\ln|\Omega_n^{b_1}|/n &- (\ln 2)b_1/2 - \epsilon(1+\epsilon)^{-1}\ln d + \xi_{\hat\alpha}\\
&> -s\ln d - s(\ln 2)/2 - s(\ln 2)/2 + \xi_{\hat\alpha}\\
&= -s(\ln d + \ln 2) + \xi_{\hat\alpha} = -s\,\ln(2d) + \xi_{\hat\alpha}\\
&> -\hat\alpha M\ln(2d) + \xi_{\hat\alpha}.
\end{aligned}
$$

Finally, since $\xi_{\hat\alpha} \ge -\hat\alpha \ln \hat\alpha$, choosing $\ln \hat\alpha < -2M \ln(2d)$, equivalently, $\hat\alpha < (2d)^{-2M}$, we gather that $\xi_{\hat\alpha} > \hat\alpha(2M)\ln(2d)$, thus,

$$-\hat\alpha M \ln(2d) + \xi_{\hat\alpha} > -\hat\alpha M \ln(2d) + 2\hat\alpha M \ln(2d) = \hat\alpha M \ln(2d).$$

At last, if we write $\kappa = \hat\alpha M \ln(2d) > 0$, then

$$\ln|\mathcal{K}_n(b_1)|/n - (1+\epsilon)^{-1}\ln|\Omega_n^{b_1}|/n - (\ln 2)b_1/2 - \epsilon(1+\epsilon)^{-1}\ln d + \xi_{\hat\alpha} > \kappa > 0.$$

Taking $b \le \kappa/\beta$ therefore assures that (4.29) holds, in turn, verifies (4.26). In summary, if we set

$$\begin{aligned} \zeta_*(\beta) &= \min\{\kappa,\, \beta\hat\alpha M,\, \beta b_2(\mathcal{G},\beta),\, 4\beta(1/2-e^{-1})/(6+d)\} \\ b_*(\beta) &= \zeta_*(\beta)/\beta, \end{aligned}$$

and choose $2\hat\alpha < b \le b_*(\beta)$ and $0 < b_1 < b - 2\hat\alpha$, then we verified (4.28). We remark that $\tau_l(\mathcal{G},\beta) \le \zeta_*(\beta)^2/(2\beta)$. This completes our proof. □

We remark that $\tau_l(\mathcal{G},\beta)$ in Proposition 2 tends to zero as $\beta \to \infty$, whereas $\tau_u(\mathcal{G},\beta)$ in Proposition 1 is uniformly bounded for all $\beta > 0$, in particular, as $\beta \to \infty$ (yet $b_2(\beta)$ tends to zero in that event). A consequence of Propositions 1 and 2 is our main result

$$\begin{aligned} \mathbf{Q}_n^\beta(J_n \notin [b_1 n, b_2 n]) &= \mathbf{Q}_n^\beta(J_n < b_1 n) + \mathbf{Q}_n^\beta(J_n > b_2 n) \\ &< e^{-\tau_l n} + e^{-\tau_u n} \le 2e^{-\tau(\mathcal{G},\beta)\, n}, \end{aligned}$$

where $\tau(\mathcal{G},\beta) = \min(\tau_u, \tau_l)$, which we restate as

Theorem **1** (Likely Interval for J_n) *Consider the number J_n of self-intersections of the weakly SAW with parameter $\beta > 0$ of length n on an infinite, locally finite connected transitive graph $\mathcal{G}$ of degree ≥ 2 and let $\mathbf{Q}_n^\beta$ denote the measure of this walk. There are positive finite constants $b_2 = b_2(\mathcal{G},\beta)$, $\hat\alpha$, $2\hat\alpha < b_*(\beta) \le b_2$ (as stated in Propositions 1 and 2) and $0 < \tau(\mathcal{G},\beta) < \infty$, all independent of n but dependent on $\mathcal{G}$ and β, such that for every $0 < b_1 < b_*(\beta) - 2\hat\alpha$ and each sufficiently large integer n,*

$$\mathbf{Q}_n^\beta(J_n \in [b_1 n, b_2 n]) > 1 - 2\,e^{-\tau(\mathcal{G},\beta)\, n}.$$

By relying on the same approach and parallel reasoning, one can prove the statement in Theorem 1 for $\tilde J_n$, the number of self-intersections at bonds: There are positive finite constants $\tilde b_2 = \tilde b_2(\mathcal{G},\beta)$, $\tilde\alpha$, $2\tilde\alpha < \tilde b_*(\beta) \le \tilde b_2$ and a constant $0 < \tilde\tau(\mathcal{G},\beta) < \infty$, all independent of n but dependent on $\mathcal{G}$ and β, such that for every $0 < \tilde b_1 < \tilde b_*(\beta) - 2\tilde\alpha$ and each sufficiently large integer n,

$$\tilde{\mathbf{Q}}_n^\beta(\tilde J_n \in [\tilde b_1 n, \tilde b_2 n]) > 1 - 2\,e^{-\tilde\tau(\mathcal{G},\beta)\, n}. \tag{4.30}$$

Observe that in the proof of Proposition 2 our strategy to introduce self-intersection events generates the same number of self-intersections at vertices as at edges of $\mathcal{G}$. Since the single repetition steps only involve one bond but two sites, certain estimates of the proof are slightly easier when considering $\tilde{J}_n$. Those results reach beyond J_n and $\tilde{J}_n$ and can be generalized to other self-intersection events of the paths of the random walk.

We conclude this section with a few remarks and open questions. Since $0 \leq J_n, \tilde{J}_n \leq n^2$, Theorem 1 and (4.30) reveal that for all large enough n,

$$\mathbf{E}_\beta J_n \in [b_1 n, b_2 n] \qquad \text{and} \qquad \tilde{\mathbf{E}}_\beta \tilde{J}_n \in [\tilde{b}_1 n, \tilde{b}_2 n]. \tag{4.31}$$

We compare these expectations to those of the simple random walk. Rewriting

$$\begin{aligned} \mathbf{E}_0 J_n &= \sum_{0 \leq i < j \leq n} \mathbf{P}_0(S_i = S_j) \\ &= \sum_{0 \leq i < j \leq n} \mathbf{P}_0(S_{j-i} = \mathbf{0}) \end{aligned} \tag{4.32}$$

and applying the local central limit theorem, one verifies the known results in $\mathbf{Z}^d$ for $d \geq 1$ that as $n \to \infty$,

$$\mathbf{E}_0 J_n = (1 + o(1)) \begin{cases} \left(\frac{2}{\pi^{1/2}}\right) n^{3/2} & d = 1, \\ \frac{1}{\pi}\, n \ln n & d = 2, \\ c_d\, n & d \geq 3 \end{cases}$$

for some positive finite constants c_d. Hence, at least in dimensions 1 and 2, we expect a significantly greater J_n of the SRW than of the weakly SAW.

What can be said about other lattices (e.g. the honey-comb lattice, the triangular lattice) and connected graphs ? How does $\mathbf{E}_0 J_n$ compare with $\mathbf{E}_\beta J_n$? What about $\tilde{J}_n$?

For instance, on the planar triangular lattice T_Δ, by writing $S_n = a_n u + b_n v$ with basis vectors $u = (0, 1)$ and $v = (\cos(2\pi/3), \sin(2\pi/3))$ in $\mathbf{R}^2$, one can apply the local central limit theorem to the random walk $\{(a_n, b_n)\}$ in $\mathbf{Z}^2$. Thus, $\mathbf{P}_0(S_i = S_j)$ approximately equals $C/|i - j|$ for some positive finite constant C, which, in view of (4.32), yields that, asymptotically for n, the expectation $\mathbf{E}_0 J_n$ is proportional to $n \ln n$ on T_Δ, too. We end with

QUESTION **4** *On graphs which are not trees, is there any order between the moments* $\mathbf{E}_\beta J_n$, $\mathbf{E}_\beta \tilde{J}_n$, $\tilde{\mathbf{E}}_\beta J_n$, *and* $\tilde{\mathbf{E}}_\beta \tilde{J}_n$ *except for the obvious ones* $\mathbf{E}_\beta \tilde{J}_n \leq \mathbf{E}_\beta J_n$ *and* $\tilde{\mathbf{E}}_\beta \tilde{J}_n \leq \tilde{\mathbf{E}}_\beta J_n$? *Which inequalities are* strict ?

References

[1] DOMB, C. AND JOYCE, G.S. (1972). Cluster expansion for a polymer chain. *J. Phys. C: Solid State Phys.* **5**, 956–976.

[2] FLORY, P.J. (1949). The configuration of real polymer chains. *Journal of Chemical Physics* **17**, 303–310.

[3] FLORY, P.J. (1974). Spatial configuration of macromolecular chains. In Markovitz, H. and Casassa, E.F. (editors), *Polymer Science: Achievements and Prospects.* Wiley Interscience, New York.

[4] HUETER, I. (2003). Proof of the conjecture that the planar self-avoiding walk has distance exponent 3/4. Preprint.

[5] MADRAS, N. AND SLADE, G. (1993). *The Self-Avoiding Walk. Probability and its Applications.* Birkhäuser Boston, Boston, MA.

IRENE HUETER, DEPARTMENT OF MATHEMATICS, BARUCH COLLEGE-CUNY, BOX B6-230, ONE BERNARD BARUCH WAY, NEW YORK, NY 10010, USA
irene_hueter@baruch.cuny.edu
http://www.math.baruch.cuny.edu/~ihueter