Delocalization of Polymers in Lower Tail Large Deviation

Riddhipratim Basu¹, Shirshendu Ganguly², Allan Sly³

¹ International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore, India. E-mail: rbasu@icts.res.in
² Department of Statistics, UC Berkeley, Berkeley, CA, USA. E-mail: sganguly@berkeley.edu
³ Department of Mathematics, Princeton University, Princeton, NJ, USA. E-mail: allansly@princeton.edu

Received: 3 May 2018 / Accepted: 23 April 2019
Published online: 13 August 2019 – © Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract: Directed last passage percolation models on the plane, where one studies the weight as well as the geometry of optimizing paths (called polymers) in a field of i.i.d. weights, are paradigm examples of models in KPZ universality class. In this article, we consider the large deviation regime, i.e., when the polymer has a much smaller (lower tail) or larger (upper tail) weight than typical. Precise asymptotics of large deviation probabilities have been obtained in a handful of the so-called exactly solvable scenarios, including the Exponential (Johansson in Commun Math Phys 209(2):437–476, 2000) and Poissonian (Deuschel and Zeitouni in Comb Probab Comput 8(03):247–263, 1999; Seppäläinen in Probab Theory Relat Fields 112(2):221–244, 1998) cases. How the geometry of the optimizing paths change under such a large deviation event was considered in Deuschel and Zeitouni (1999) where it was shown that the paths [from (0, 0) to (n, n), say] remain concentrated around the straight line joining the end points in the upper tail large deviation regime, but the corresponding question in the lower tail was left open. We establish a contrasting behaviour in the lower tail large deviation regime, showing that conditioned on the latter, in both the models, the optimizing paths are not concentrated around any deterministic curve. Our argument does not use any ingredient from integrable probability, and hence can be extended to other planar last passage percolation models under fairly mild conditions; and also to other non-integrable settings such as last passage percolation in higher dimensions.

1. Introduction and Main Results

Last passage percolation models on the plane are paradigm examples of models believed to be in the KPZ universality class. In these models, vertices of \( \mathbb{Z}^2 \) are equipped with independent and identically distributed random weights. The weight of a path is the sum of the weights along it, and the last passage time between two points is obtained by maximizing the weight among all directed paths between them (see Sect. 1.1 for precise definitions). Although the asymptotic behaviour is believed to be universal under mild
conditions on the passage time distribution, detailed understanding of these models has so far been mostly restricted to a handful of exactly solvable cases where very fine information, both algebraic and geometric, is obtained using formulae from integrable probability. Although our results hold for last passage percolation with fairly general edge weight distribution; for the sake of concreteness we shall focus, for much of this article, on the interesting special case of the exactly solvable model with exponentially distributed vertex weights, while deferring until later the extension to more general settings.

This case of Exponential directed last passage percolation (LPP) is very well studied, in particular because of the correspondence with Totally Asymmetric Simple Exclusion Process (TASEP) on a line. Using the understanding of invariant measures of TASEP, already in 1981, Rost [16] evaluated the limiting shape for this model; in particular he showed the following. Let $L_n$ denote the last passage time from $(0,0)$ to $(n,n)$, then $E L_n / n \to 4$. Rost’s results [16] in particular show that the limit shape for exponential last passage percolation is strictly concave; this, together with some basic concentration estimates (e.g. in [19]), imply that the maximal path (henceforth called the geodesic, or the polymer) from $(0,0)$ to $(n,n)$ is with high probability concentrated around the straight line joining the two points. More precise results were obtained later using exact determinantal formulae: Johansson [11] established the $n^{1/3}$ fluctuation of $L_n$ and a Tracy-Widom scaling limit; a more precise version of that also implies that the fluctuation of the geodesic around the diagonal line is of the order $n^{2/3}$ (see [5,12] for more details).

Along with the typical behaviour; large deviation behaviour of $L_n$ (i.e., when the deviation of $L_n$ from $4n$ is linear in $n$) has also attracted attention. In fact Johansson [11] obtained large deviation rate functions for $L_n$; i.e., the precise rates of decay for probabilities that $L_n$ is either much larger or much smaller than typical (see Theorem 1.4). There has been a great deal of interest in the general theory of large deviations to understand the geometric consequences of conditioning on rare large deviation events. In this paper we study the geometry of the geodesic in last passage percolation when the passage time $L_n$ is conditioned to be atypical.

It is at least heuristically not too hard to see that, in the upper tail large deviations regime, i.e., when the last passage time is conditioned to be macroscopically larger than typical; the geodesic is still localized with high probability around the diagonal. The picture in the lower tail large deviations regime is more complicated. We establish a contrasting delocalization result in this case. Our main result, Theorem 1, shows that conditioned on the last passage time being much smaller than typical, the geodesic is not localized around any deterministic curve. The different behaviour in the two tails is intimately connected to the different speeds at which large deviation occurs in the upper tail and lower tail regimes respectively, (see Sect. 1.3 for further elaboration along these lines).

In the context of Poissonian directed last passage percolation (henceforth to be referred to as Poissonian LPP) on the plane, another model in the KPZ universality class, which has the same qualitative behaviour as the exponential model, this question was investigated by Deuschel and Zeitouni [9], who obtained explicit formulae for the rate function and showed that conditioned on the upper tail large deviation event, the geodesic is indeed localized around the diagonal. However the question about whether or not similar behaviour is observed in the lower tail large deviation regime was left open. We answer this question by showing a similar delocalization behaviour in the lower tail in this case as well (see Theorem 2).
Although both Exponential and Poissonian LPP models are exactly solvable, our arguments do not use integrability in any crucial way, and hence can be extended to more general settings. We prove a similar delocalization result in the lower tail large deviation result of last passage percolation on $\mathbb{Z}^2$ with some mild condition on the passage time distribution (see Theorem 3 for the precise conditions). We also establish an extension to higher dimensional LPP models (see Theorem 4). As far as we are aware these are the first results on the geodesic geometry in such a general setting.

Even though our results do not use integrability, the special case of Exponential LPP leads to certain simplifications making the basic argument more transparent. We shall therefore, to start with, restrict ourselves to this case, while postponing the discussion about the other cases. We now move towards precise model definitions and statement of the main result in the Exponential case.

1.1. Model definitions. Let $\Pi = \{X_v: v \in \mathbb{Z}^2\}$ be a collection of i.i.d. Exponential random variables with rate 1. Consider the following partial order $\preceq$ on $\mathbb{Z}^2$: we say $(x_1, y_1) \preceq (x_2, y_2)$ if $x_1 \leq x_2$ and $y_1 \leq y_2$. For $u \preceq v$, a directed path $\gamma$ from $u$ to $v$ is defined as an up/right path starting at $u$ and ending at $v$, i.e., $\gamma := \{u = u_0 \preceq u_1 \preceq \cdots \preceq u_k = v\}$ is a path in $\mathbb{Z}^2$ where for each $i$, $u_i - u_{i-1} = (1, 0)$ or $(0, 1)$. For a directed path $\gamma$ as above, the passage time (or, as we shall often say, length) of $\gamma$, denoted $L(\gamma)$ is defined as $\sum_{i=0}^{k} X_{u_i}$.

Definition 1.1. For $u \preceq v \in \mathbb{Z}^2$, define the last passage time $L(u, v)$ from $u$ to $v$ by

$$L(u, v) := \max_{\gamma} L(\gamma)$$

where the maximum is taken over all directed paths from $u$ to $v$. The maximizing path will be called the geodesic between $u$ and $v$ which will be denoted by $\Gamma(u, v)$.\(^1\)

Let us now introduce some notations. For $x, y \in \mathbb{Z}_+$, we shall denote by $L_{x,y}$ the last passage time from $0 := (0, 0)$ to $(x, y)$. When $x = y(= n, \text{say})$ we shall simplify the notation even further and denote the last passage time by $L_n$. The geodesic from $0$ to $n := (n, n)$ shall be denoted $\Gamma_n$. Let $\gamma : [0, 1] \to [0, 1]$ denote a continuous increasing function with $\gamma(0) = 0$ and $\gamma(1) = 1$. We define the $(\varepsilon, n)$-cylinder neighbourhood of $\gamma$, denoted $\gamma_n^\varepsilon$ by

$$\gamma_n^\varepsilon = \left\{(x, y) \in [0, n]^2 : |y - n\gamma(n^{-1}x)| \leq \varepsilon n\right\};$$

i.e., it is a cylinder of width $\varepsilon n$ around the path from $0$ to $n$ that is obtained by scaling up $\gamma$. The following result is standard (see e.g. [4] for a much stronger result).

Theorem 1.2. Let $\mathbb{I}$ denote the identity function on $[0, 1]$, and let $\varepsilon > 0$ be fixed. Let $E_n$ denote the event that $\Gamma_n$ is contained in $\mathbb{I}_n^\varepsilon$. Then $\mathbb{P}(E_n) \to 1$ as $n \to \infty$.

Observe that this result asserts that the geodesic $\Gamma_n$ with high probability has Hausdorff distance $o(n)$ from the diagonal line joining $0$ and $n$. This is a very general result and essentially uses only the strict concavity of the limit shape. For exactly solvable models of last passage percolation, quantitatively optimal stronger variants of this result

\(^1\) Observe that there is almost surely a unique maximizing path between any two vertices, by continuity of Exponential random variables.
Theorem 1 says that on the lower tail large deviation event, the geodesic $\Gamma_n$ is unlikely to be contained in such a cylinder for any fixed curve $\gamma$. See Sect. 1.2 for more elaboration on this. As mentioned in the introduction, our main result in this paper shows that this behaviour changes in the lower tail large deviations regime. Formally, for $\delta \in (0, 4)$ let $L_\delta$ denote the event $L_n \leq (4 - \delta)n$. We have the following theorem.

**Theorem 1.** For each $\delta \in (0, 4)$, and an increasing continuous function, $\gamma : [0, 1] \to [0, 1]$, with $\gamma(0) = 0$ and $\gamma(1) = 1$, for any $\varepsilon > 0$, there exists $\varepsilon' > 0$, such that for all large enough $n$,

$$\mathbb{P}(\Gamma_n \subseteq \gamma_{\varepsilon'}^n | L_\delta) \leq \varepsilon.$$

Thus Theorem 1 asserts that there does not exist any deterministic curve $\gamma$, such that $\Gamma_n$ is localized around that curve with high probability (see Fig. 1 for an illustration of the theorem). One however believes the much stronger following statement to be true: that in fact, given any anti-diagonal, (say $\{x + y = n\}$), the location where the geodesic intersects it (i.e., vertex $\Gamma_n \cap \{x + y = n\}$) should also not be concentrated on any set of size $o(n)$. We will elaborate further on this point and the associated difficulties later in the article at the end of Sect. 3.

In the next Sect. we continue our discussion regarding typical and rare behaviour of polymer models and variants of Theorem 1, in other related settings including a class of non-integrable models, and compare such results with existing literature.

### 1.2. Background, previous works and our contributions.

For polymer models in the KPZ universality class (last passage percolation is a general example of a zero temperature polymer model), the geometric properties of polymers (i.e., the geodesics in our context) has been an object of fundamental study. The three scaling exponents $(1, 1/3, 2/3)$ are characteristic of the KPZ universality class that corresponds to polymer length, length fluctuation and spatial decay of correlation. In the context of last passage percolation it can be illustrated as follows. For the geodesic from $\mathbf{0}$ to $\mathbf{n}$, the length of the geodesic is of
the order $n^1$, the fluctuation of the length is of the order $n^{1/3}$ and the typical distance of the geodesic to the straight line joining the two points is of the order $n^{2/3}$. As mentioned before, this behaviour is expected to be universal for last passage percolation under mild condition on the passage times, but is rigorously known only for a handful of models including last passage percolation on $\mathbb{Z}^2$ with exponential and geometric weights, and also Poissonian LPP on $\mathbb{R}^2$ where, one maximizes the number of points on an oriented path in a field of Poisson points on the plane.

Length fluctuations of the order $n^{1/3}$ was first proved for the Poissonian LPP in the seminal work of Baik, Deift and Johansson [1], who also proved the weak convergence to GUE Tracy-Widom distribution after suitable centering and scaling. The corresponding result in Exponential last passage percolation is due to Johansson [11]. For completeness, let us recall the standard results in this case. Recall the last passage time $L_{x,y}$ from $(0,0)$ to $(x,y)$. The first order behaviour for $L_{x,y}$ was established in [16].

**Theorem 1.3.** Let $x, y > 0$ be fixed real numbers. Then

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E} L_{[nx],[ny]} = G(x, y) = (\sqrt{x} + \sqrt{y})^2.$$  

In particular, for $x = y = 1$, this shows that $\mathbb{E} L_n = (4 + o(1)) n$. Moreover, for $x, y > 0$, it is known that $n^{-1/3} (L_{[nx],[ny]} - nG(x,y))$ converges weakly. It is however, much easier to show (and the argument is much more general) concentration of $L_{[nx],[ny]}$ at scale $\sqrt{n}$ (see for example, [19]). Also important is the boundary of the limit shape $\{(x,y) \in \mathbb{R}^2 : G(x,y) = 1\}$ which can be observed to be strictly concave. This implies that the growth of $L_{x,y}$ is fastest in the diagonal direction $(1,1)$. This, together with the above concentration result establishes that $\Gamma_n$ is concentrated around the straight line joining $0$ and $n$ as already stated in Theorem 1.2 that was established in [8] in the context of Poissonian LPP and its variants. Sharp exponent of transversal fluctuation (i.e., the maximum vertical distance between a geodesic from $0$ to $n$ and the diagonal line joining two points) was obtained in [12] which showed that the maximum transversal fluctuation in $n^{2/3+o(1)}$ with high probability in the Poissonian case. The same holds in the Exponential case; see [5, Theorem 11.1] for the statement of a quantitatively sharper result.

Work on large deviations in polymer models goes back to Kesten [13] in 1986 who considered large deviation problems in the related setting of first passage percolation. For $\delta > 0$, let $\mathcal{U}_\delta = \{ L_n \geq (4+\delta)n \}$ denote the upper tail event analogous to $\mathcal{L}_\delta$ already defined in the statement of Theorem 1. A straightforward adaptation of the argument of [13] shows that the log probabilities for the upper tail event scales as $n$, whereas the log probabilities scale as $n^2$ for the lower tail. In [11] a precise rate function was established:

**Theorem 1.4.** ([11]): There exist functions $I_u(\delta), I_l(\delta)$ such that $I_u(\delta) \in (0, \infty)$ for all $\delta > 0$ and $I_l(\delta) \in (0, \infty)$ for $\delta \in (0,4)$ such that

$$\lim_{n \to \infty} \frac{1}{n} \log(\mathbb{P}(L_n \geq (4+\delta)n)) = -I_u(\delta)$$

$$\lim_{n \to \infty} \frac{1}{n^2} \log(\mathbb{P}(L_n \leq (4-\delta)n)) = -I_l(\delta).$$

Existence of the rate function for the upper tail follows from a standard sub-additive argument, and using appropriate concentration estimates one can show non-triviality of the rate function. However the proof of the lower tail result depends on the exact
determinantal formulae for this specific model. Johansson provides an explicit formula for the upper tail rate function, whereas the lower tail rate function was not evaluated explicitly.

The analogue of Theorem 1.4 was proved for Poissonian LPP in [9, 18] using connections to longest increasing subsequences for permutations and the RSK correspondence to Young Tableaux and connections to Hammersley’s process. Specifically the right tail behaviour was established in [18] without using the Young Tableaux combinatorics, and a similar analysis for the right tail large deviation for Exponential LPP was done in [17] using connections to TASEP. In [9], the geometry of geodesics in the large deviation regime has been investigated in this setting. Conditioned on the upper tail large deviation event, it was shown that the geodesics remain localized around the diagonal. Their method of proof can be adapted to our setting of the Exponential LPP in a straightforward manner to show that in the notation of Theorem 1.2, except for an event of exponentially (in $n$) small conditional probability given $U_\delta$, we have $\Gamma_n \subseteq I_\varepsilon n$.

The analysis of the lower tail rate function in [9] involves solving a variational problem for shapes of Young Tableaux and did not provide any geometric information about the optimal path. In particular, whether or not the path is localized around a deterministic curve conditioned on the lower tail large deviation event was mentioned as Open Problem 2 in [9]. Our arguments proving Theorem 1 can be adapted in this setting too to show that conditioned on the lower tail large deviation event the path is not localized around any deterministic curve.

**Theorem 2.** Fixing $\delta$, for any increasing continuous $\gamma : [0, 1] \to [0, 1]$ with $\gamma(0) = 0$ and $\gamma(1) = 1$, there exist $\varepsilon > 0$, such that

$$\mathbb{P}(E_{\gamma, n} | L_\delta) \to 1$$

as $n \to \infty$, where $E_{\gamma, n}$ denotes the event that there exists a geodesic $\Gamma_n$ between $0$ and $n$ that is not contained in $\gamma_n^\varepsilon$, where $\gamma_n^\varepsilon$ is defined as in (1) replacing $[0, n]^2$ by $[0, n]^2$.

Note the subtle qualitative difference between Theorems 1 and 2. In the former, by continuity of the exponential variables, the geodesic $\Gamma_n$ is well defined, whereas for the latter, due to the discrete nature of the setting, there can be several geodesics between points $0$ and $n$. Hence whereas the former says that conditioned on $L_\delta$, with high probability the geodesic does not lie within a narrow sausage around any deterministic function, the latter theorem shows existence of geodesics outside such a sausage, while not precluding existence of some geodesic within the sausage.

Thus Theorems 1 and 2 both exhibit a transition from localization to delocalization in going from the upper tail to the lower tail of the large deviation regime. The fundamental idea behind the proof of Theorem 2 is the same as Theorem 1; however the proof of the former involves certain additional technical steps including a discretization argument, which, in particular leads to the subtly different statement of the result in the two cases. For the sake of brevity and maintaining transparency, we have chosen to not include the details of the proof of Theorem 2 in this article. A precise statement together with all details can be found in [3], a longer version of this paper on arXiv. However we do mention a key difference between the two cases briefly in Sect. 1.3.3, see also Fig. 2b.

Finally we emphasize that even though the above mentioned theorems are in the integrable setting of Exponential LPP or Poissonian LPP on $\mathbb{Z}^2$ and $\mathbb{R}^2$ respectively, our argument actually is far more general. In Sect. 4 we consider LPP on $\mathbb{Z}^2$ with general passage times which makes the model lose its integrable structure. Under some smoothness conditions, e.g., monotonicity or log-concavity, on the density of the passage
To prove strong concentration, we expose the environment except for one anti-diagonal $D$. The conditional law of $X_v$ for $v \in D$, given the rest of the the vertices, is an Exponential random variable conditioned to be less than $M_v$ where $M_v = (4 - \delta)n - |\gamma_v|$ where $\gamma_v$ is the best path passing through $v$. Moreover $X_v$ for $v \in D$ are independent of each other. Thus $(4 - \delta)n - L_n \leq \inf_{v \in D} M_v - X_v$ and the latter quantity can be seen to be $O(1/n)$ if $M_v < M$ for a fixed constant $M$ for $\Theta(n)$ many $v$’s. The same proof strategy works for Poissonian LPP; however in this case we expose the point process everywhere except inside small boxes along the anti-diagonal.

While on the topic of non-integrable models, we should mention that the absence of exact formulas present significant challenges towards just showing the existence of a large deviation rate function analogous to Theorem 1.4. Beyond the exactly solvable regime, such a result was, until recently, known only in a single setting with special geometric constraint [6]. In [2], we prove the existence of the rate function in the context of general first and last passage percolation models. However, our results in this paper do not require the existence of the rate function and only relies on the fact that the speed of decay for the lower tail large deviation probability is $n^2$.

### 1.3. Outline of the Proof

We present the key ideas in the proof of Theorem 1 in this subsection. We start with the observation that the speed $n^2$ for the log-probability of $\mathcal{L}_\delta$ in Theorem 1.4 follows by noting that given $\delta$ there exists $C > 0$, such that there are $\Theta(n)$ many disjoint translates of the strip of width $C$ around the main diagonal joining $0$ and $n$ such that the maximum weight increasing path in each of these strips typically has length at least $(4 - \delta)n$. This implies that to achieve the event $\mathcal{L}_\delta$, one needs all of those paths to have values which is smaller than typical. This forces $\Theta(n^2)$ many vertices to have weights smaller than typical and hence such an event would be exponential in $n^2$ unlikely. This is the only aspect of the large deviation event that our proof relies on.
(hence allows us to prove theorems in the non-integrable setting as well). All of the above is made precise in the proof of Proposition 2.1 where we show that while in the typical environment we have by law of large numbers \( \sum_{0 \leq v \leq n} X_v = n^2 + O(n) \) with high probability, in fact conditioned on \( L_\delta \), with high probability
\[
\sum_{0 \leq v \leq n} X_v \leq (1 - c)n^2
\] 
for some \( c = c(\delta) > 0 \). Using the above the proof of Theorem 1 has broadly two parts:

- **Strong concentration of the length of conditional geodesic:** First we show that conditional on \( \{L_n \leq (4 - \delta)n\} \), with high probability \( L_n \) is concentrated around \( (4 - \delta)n \) at scale \( \frac{1}{n} \). (We elaborate why this is true in Sect. 1.3.1)
- **Anti-concentration of constrained paths in the conditional environment:** We shall show that, conditional on \( L_\delta \), the best path that is restricted to stay within a deterministic set of vertices of size \( \epsilon n^2 \), is concentrated around \( (4 - \delta)n \) at scale \( \frac{1}{n} \), only with probability that decays to zero with \( \epsilon \).

Combining the above, Theorem 1 follows easily. We now elaborate further on the above points.

1.3.1. Strong concentration of conditional geodesic. We have the following theorem.

**Theorem 1.5.** Fix any \( \delta \in (0, 4) \). Given any \( \epsilon > 0 \) there exists \( H > 0 \) such that
\[
\mathbb{P} \left( L_n \geq (4 - \delta)n - \frac{H}{n} \mid L_\delta \right) \geq 1 - \epsilon.
\]

To see why this should be true, note that for any anti-diagonal \( D \) i.e., a set of the form \( \{(x, y) : x + y = k\} \), (see Fig. 2 where the anti-diagonal for \( k = n \) is highlighted) the typical value of \( \sum_{v \in D} X_v \) is \( |D| + O(\sqrt{|D|}) \), since exponential variables have mean one and the sum is well concentrated. However (2) implies that, conditional on \( L_\delta \), typically there exists \( \Theta(n) \) many anti-diagonals \( D \) (each with size linear in \( n \)) such that \( \sum_{v \in D} X_v \leq (1 - c)|D| \) for some \( c = c(\delta) > 0 \). At this point we use the following observation: fixing any anti-diagonal \( D \), if we condition on the weights of the remaining vertices \( \{X_v : v \notin D\} \) (the green region in Fig. 2a), and the event \( L_\delta \), then conditionally, \( \{X_v : v \in D\} \) (the weights on \( D \)) is nothing but a collection of independent variables where \( X_v \) follows the law of standard exponential variable conditioned to be less than \( R_v \) which is a deterministic function of the weights of the vertices in the green region. More precisely, \( R_v = (4 - \delta)n - M_v \) where
\[
M_v = \max_{\gamma} |\gamma| - X_v,
\]
where the maximum is taken over all directed paths \( \gamma \) from \( 0 \) to \( n \) that pass through \( v \). Note that any directed path from \( 0 \) to \( n \) intersects any anti-diagonal at exactly one point. Thus \( M_v \) (and hence \( R_v \)) is indeed a deterministic function of all the variables \( \{X_v : v \notin D\} \). At this point we shall conclude that for anti-diagonals \( D \) such that \( \sum_{v \in D} X_v \leq (1 - c)|D| \), with high probability \( R_v \) must be uniformly bounded by some \( M \) for at least \( \Theta(n) \) many vertices in \( D \).

It might be useful to think about this in the following way. For each \( v \) on an anti-diagonal, the event \( L_\delta \) forces a ‘barrier’ \( R_v \), and \( X_v \) is an Exponential random variable conditioned not to exceed this barrier. Clearly the \( v \) such that \( X_v \) gets closest to its barrier...
$R_v$ will be the unique point on the anti-diagonal that the maximum length path passes through and consequently $(4 - \delta)n - L_n$ is $R_v - X_v$. Now let us pretend that all the barriers are $M$ (this is not true but the calculation is qualitatively the same as long as $R_v \leq M$ for a linear number of $v \in D$), and the proof is complete by observing that $M - \max_{1 \leq i \leq n} Y_i = O\left(\frac{1}{n}\right)$ with high probability where $Y_i$ are i.i.d. Exp(1) variables conditioned to be at most $M$.

1.3.2. Anti-concentration of restricted paths in the conditional environment. To make a formal statement we need the following notations. For any $A \subseteq \mathbb{Z}^2$, let $\Gamma_n(A)$ (resp. $L_n(A)$) be the longest directed path (resp. length) which lies entirely in $A$. We then have the following theorem.

**Theorem 1.6.** Fix any $\delta > 0$. Then given any $H$ and $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that for every deterministic set $A \subseteq [0, n]^2$, with $|A| \leq \varepsilon' n^2$ we have

$$\mathbb{P}\left( L_n(A) \geq (4 - \delta)n - \frac{H}{n} | \mathcal{L}_\delta \right) \leq \varepsilon.$$

Let us attempt to give an informal reasoning for the above theorem. Clearly $L_n(A)$ only depends on $\{X_v : v \in A\}$. We can decompose this data into two parts: $S = \sum_{v \in A} X_v$ and $\{\frac{X_v}{S} : v \in A\}$ which are independent by the properties of Exponential random variables. It is well known that $S$ is distributed as a Gamma($|A|$) random variable. Now conditioned on $\{X_v / S : v \in A\}$ and the field outside $A$ and $\mathcal{L}_\delta$ (obviously all the pieces of the data have to be compatible with $\mathcal{L}_\delta$ for this to make sense), by similar arguments as in Sect. 1.3.1, the conditional distribution of $S$ is the distribution of a Gamma($|A|$) variable conditioned to be less than some value $S_0$ where $S_0$ is a measurable function of the sigma field being conditioned on and is $O(n^2)$ for most of the realizations of the data. It then follows that the typical value of $S$ under the above conditioning is around $(1 - \frac{1}{|A|})M$. As argued before as well, $(4 - \delta) - L_n(A)$ is governed by how close $S$ is to $M$; and in particular the event $L_n(A) \geq (4 - \delta)n - \frac{H}{n}$ would require $S > (1 - \frac{C}{n^2})M$, for some $C$ which just depends on $\delta$ and $H$. Thus taking $A \leq \varepsilon' n^2$ for some small enough $\varepsilon'$ will complete the proof.

The main work of this paper goes into proving Theorems 1.5 and 1.6. We shall, however, first provide the immediate proof of Theorem 1 using these.

**Proof of Theorem 1.** Fix $\delta \in (0, 4)$ and $\varepsilon > 0$. Let $H > 0$ be such that the conclusion of Theorem 1.5 holds for this choice of $\delta$ and $\varepsilon$, and let $\varepsilon'$ be such that the conclusion of Theorem 1.6 holds for this choice of $\delta$, $\varepsilon$ and $H$. Observe that the set $\gamma^{\varepsilon'}_n$ [see (1)] has size $O(\varepsilon' n^2)$. Thus choosing the set $A$ in Theorem 1.6 to be $\gamma^{\varepsilon'}_n$, it follows that

\[
\mathbb{P}(\Gamma_n \subseteq A \mid \mathcal{L}_\delta) = \mathbb{P}(\Gamma_n = \Gamma_n(A) \mid \mathcal{L}_\delta) \\
\leq \mathbb{P}(L_n \leq (4 - \delta)n - \frac{H}{n} \mid \mathcal{L}_\delta) + \mathbb{P}(L_n(A) \geq (4 - \delta)n - \frac{H}{n} \mid \mathcal{L}_\delta) \\
\leq \varepsilon + \varepsilon.
\]

\[\square\]

We remark that the choice of the definition of the $(\varepsilon, n)$ cylinder in (1) is not canonical and similarly one can also take the paths $\gamma$ in the statement of Theorem 1, to be directed paths instead of graphs of honest functions. Minor variants of the above arguments would yield the same result in such situations and we omit the details for brevity.
1.3.3. Generalization to other settings.

(1) In the case of Poissonian LPP, due to the discrete nature of the problem we have extreme concentration, i.e., with high probability the length of the geodesic equals the largest integer smaller than the barrier value of \((2 - \delta)n\) forced by the event \(\mathcal{L}_\delta\). This is achieved using the same strategy as above, however instead of anti-diagonals of vertices, we now analyze anti-diagonals formed by small boxes (see Fig. 2b).

(2) A close inspection at the above proof technique for anti-concentration reveals that the reliance on the properties of Gamma distribution is not crucial. In fact in the proof of Theorem 3, we prove anti-concentration under pretty general distributional assumptions where \(S\) does not follow a nice distribution. Note however that the above anti-concentration result is only expected if the distribution of \(S\) has some smoothness. This is the reason why in the setting of Poissonian LPP we do not expect such a result to hold leading to a difference in the nature of Theorems 1 and 2 (see the discussion following the statement of the latter).

1.4. Organization of the paper. The rest of this paper is organized as follows. In Sect. 2 we prove Theorem 1.5. We prove the anti-concentration estimate Theorem 1.6 in Sect. 3. In Sect. 4, we prove delocalization for general last passage percolation models on the plane, going beyond the exactly solvable regime. We finish with discussions of higher dimensional extensions in Sect. 4.2.

2. Strong Concentration of the Conditional Geodesic

In this section we prove Theorem 1.5. The crucial first step is to obtain a geometric interpretation of the \(n^2\) speed for the LDP for the lower tail, showing that under the conditioning the sum of all edge weights becomes macroscopically smaller.

**Proposition 2.1.** Given \(\delta > 0\) there exists \(\varepsilon \in (0, \frac{1}{16})\) such that

\[
P \left( \sum_{v \in [0,n]^2} X_v \leq (1 - 4\varepsilon)n^2 \mid \mathcal{L}_\delta \right) \geq 1 - e^{-cn},
\]

for some \(c = c(\delta)\) and for all \(n\) large enough. (We write \(4\varepsilon\) instead of \(\varepsilon\) to avoid notational cluttering later.)

Proof of Proposition 2.1 will depend on the FKG inequality, and the argument sketched in the beginning of Sect. 1.3 showing that the large deviation occurs at speed \(n^2\). We first need to set up some notation. For \(i \in \mathbb{Z}\) and \(K \in \mathbb{N}\), let Strip\(_i^K\) be defined as follows.

\[
\text{Strip}\_i^K = \{(x, y) \in [0, n]^2 : |x - y - 4iK| \leq K\};
\]

that is, Strip\(_i^0\) is the strip of width \(2K\) around the main diagonal \(\{x = y\}\), and the other strips are its translates by integer multiples of \(4K\). Note that the above definition ensures that Strip\(_i^K\) are disjoint for different values of \(i \in \mathbb{Z}\). The next lemma shows that with high probability, in a typical environment, for any strip that is not too far from the main diagonal, there will be one path through each strip with length close to that of the longest path. Recall the notation \(L_n(A)\) from the statement of Theorem 1.6. We have the following lemma.
Lemma 2.2. Fix $4 > \delta > 0$. There exists $K, c_0, c' > 0$ depending on $\delta$ such that with probability at least $1 - e^{-c'n}$ simultaneously for all $i \in \mathbb{N}$ with $i \leq c_0n$ we have

$$L_n(\text{Strip}_i^K) \geq (4 - \frac{\delta}{2})n.$$  

We first prove the following preparatory lemma.

Lemma 2.3. Fix $\delta \in (0, 4)$. There exists $K$ sufficiently large and $c' > 0$ so that for all $n$ sufficiently large we have

$$\Pr \left( L_n(\text{Strip}_0^K) \geq (4 - \frac{\delta}{4})n \right) \geq 1 - e^{-c'n}.$$  

Proof. Using Theorem 1.3, choose $K$ sufficiently large so that $\mathbb{E}L_K \geq (4 - \frac{\delta}{2})K$. Assume without loss of generality that $n$ is a multiple of $K$. Now we use super-additivity to argue that

$$L_n(\text{Strip}_0^K) \geq L(0, K) + L(K, 2K) + \ldots + L(n - K, n),$$

where $L(\cdot, \cdot)$ was defined in Definition 1.1. Now note that all the terms on the right hand side above are i.i.d. random variables distributed as $L_K$, and hence the proof follows from standard exponentially small probability bounds for deviation below the mean for sums of such variables. \hfill \Box

We are now ready to prove Lemma 2.2.

Proof of Lemma 2.2. For $i \geq 0$, let $L_i^K$ denote the length of the best path from $(4iK, 0)$ to $(n, n - 4iK)$ that is contained in $\text{Strip}_i^K$. Observe that by translation invariance, for all $n$ sufficiently large, $L_i^K$ is distributed as $L_{n-4i|K}$, (see Fig. 3). Fixing $\delta$, choose $K$ sufficiently large depending on $\delta$ so that the conclusion of Lemma 2.3 holds. Now choose $c_0$ sufficiently small so that $(4 - \frac{\delta}{2})(1 - 4c_0K) \geq (4 - \frac{\delta}{2})$. The proof is now completed by using Lemma 2.3 and taking a union bound over all $i \in \mathbb{N}$ with $i \leq c_0n$. \hfill \Box

Fig. 3. Disjoint strips of constant width have disjoint paths of weight at least $(4 - \delta/2)n$ and hence the total weight in each of these strips in the large deviation noise field must be at least $\frac{\delta}{2}n$ smaller than the typical counterpart.
We state another simple lemma without proof, which is a straightforward consequence of FKG inequality. Recall definition of $\Pi$ from Sect. 1.1. Let $\Pi^* = \{X^*_v\}_{v \in \mathbb{Z}^2}$ be $\Pi$ distributed conditionally on $L_\delta$.

**Lemma 2.4.** There is a coupling $(\Pi^*, \Pi)$ such that almost surely, $X^*_v \leq X_v$ for all $v \in \mathbb{Z}^2$.

We are finally ready to prove Proposition 2.1.

**Proof of Proposition 2.1.** Consider the coupling $(\Pi^*, \Pi)$ from the previous lemma. For all $0 \leq i \leq c_0 n$, (where $c_0$ appears in Lemma 2.2) note that

$$\sum_{v \in \text{Strip}_i^K} (X_v - X^*_v) \geq L_{i,n} - L_{i,n}^* \geq \frac{\delta}{2} n$$

with probability at least $1 - e^{-cn}$, where $L_{i,n}^*$ denotes the natural analogue of $L_{i,n}$ corresponding to $\Pi^*$ and the last probability bound follows from Lemma 2.3 and that by definition $L_{i,n}^* \leq (4 - \delta)n$. The proof is now complete by summing the above inequality over $0 \leq i \leq c_0 n$ and noting that

$$\sum_{v \in \llbracket 0, n \rrbracket^2} (X_v - X^*_v) \geq c_0 n \sum_{i=0}^{c_0 n} \left( \sum_{v \in \text{Strip}_i^K} (X_v - X^*_v) \right) \geq \frac{\delta c_0 n^2}{2}$$

with probability at least $1 - e^{-cn}$, where the last inequality follows from (4) and union bound over $0 \leq i \leq c_0 n$. The proof is now complete by choosing $4\varepsilon = \frac{\delta c_0}{4}$, and using the straightforward consequence of concentration of sum of exponential random variables, that, $\sum_{v \in \llbracket 0, n \rrbracket^2} X_v \leq (1 + \frac{\delta c_0}{4})n^2$ with probability at least $1 - e^{-cn^2}$ for some $c = c(\delta) > 0$. $\square$

Before continuing with the proof of Theorem 1.5, we need to set up some more notation.

**For the remainder of this section $\delta \in (0, 4)$ and $\varepsilon$ will be fixed such that the conclusion of Proposition 2.1 holds.**

For $i \in [0, 2n]$, let $D_i$ denote the $i$-th anti-diagonal, i.e.,

$$D_i = \{v = (x, y) \in [0, n]^2 : x + y = i\}$$

and let,

$$I = \{i \in \mathbb{N} : |i - n| \leq (1 - \sqrt{\varepsilon})n\}.$$

We shall restrict our attention to those $D_i$’s with $i \in I$, in particular these $D_i$’s all have size linear in $n$. For $i \in I$, let $\mathcal{F}_i$ denote the sigma algebra generated by $\{X_v : v \notin D_i\}$. Let $\gamma_v$ denote the longest directed path from $0$ to $n$ that passes through $v$ (as already mentioned any such directed path intersects $D_i$ exactly once). Observe that $\gamma_v$ is $\mathcal{F}_i$ measurable, and so is $L(\gamma_v) - X_v$. For any $v \in D_i$, set

$$R_v = (4 - \delta)n - (L(\gamma_v) - X_v).$$

Notice that $R_v$ is $\mathcal{F}_i$ measurable. We now have the following easy lemma.
**Lemma 2.5.** Fix $i \in I$ and condition on $\mathcal{F}_i$, so that $\mathcal{F}_i$ is compatible with $\mathcal{L}_\delta$ (this implies $R_v > 0$ for each $v \in D_i$). Then conditional on $\mathcal{F}_i$, and the event $\mathcal{L}_\delta$, the random variables $\{X_v : v \in D_i\}$ are independent and the conditional distribution of $X_v$ is given by that of an exponential random variable with rate one conditioned to be at most $R_v$ (we denote this law by $\text{Exp}(0, R_v)$).

**Proof.** The proof follows by observing that the only effect of conditioning on $\mathcal{L}_\delta$ and $\mathcal{F}_i$, is that, for any $v \in D_i$, $X_v$ is restricted to be less than $R_v$ since otherwise the conditioning imposed by $\mathcal{L}_\delta$ would be violated. □

The next lemma is an easy consequence of properties of truncated exponential variables.

**Lemma 2.6.** Let $X_1, X_2, \ldots, X_m$ be independent with $X_i \sim \text{Exp}(0, f_i)$ where $f_i > 0$ for all $i \in [1, m]$. Let $M > 0$ be such that $\mathbb{E}(\text{Exp}(0, M)) = (1 - \varepsilon/4)$. Suppose $\#\{i : f_i \leq M\} \leq \varepsilon m/2$. Then there exists $c > 0$ such that

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i \leq (1 - \varepsilon)m\right) \leq e^{-cm}.$$  

**Proof.** The proof follows from observing that $\text{Exp}(0, M)$ is stochastically increasing in $M$ and hence letting $A = \{i : f_i > M\}$ we see that $\mathbb{P}\left(\sum_{i=1}^{m} X_i \leq (1 - \varepsilon)m\right) \leq \mathbb{P}\left(\sum_{i \in A} Y_i \leq (1 - \varepsilon)m\right)$ where $Y_i$ are i.i.d. $\text{Exp}(0, M)$ variables. The proof now follows from exponential concentration of sums of i.i.d. variables. □

Let $M$ as in the above lemma be fixed for the remainder of this section. For $i \in I$, let us define the event

$$\mathcal{M}_i := \left\{\#\{v \in D_i : R_v \leq M\} \geq \varepsilon |D_i|/2\right\}.$$  

Observe that $\mathcal{M}_i$ is $\mathcal{F}_i$ measurable. Also let us denote by $C_H$, the event from the statement of Theorem 1.5, i.e.,

$$C_H := \left\{L_n \geq (4 - \delta)n - \frac{H}{n}\right\}.$$  

We want to prove that $C_H$ happens with large probability conditional on $\mathcal{L}_\delta$. The following lemma is the first step in this direction that demonstrates the usefulness of the events $\mathcal{M}_i$.

**Lemma 2.7.** For any $\varepsilon_1 > 0$, there exists $H > 0$ sufficiently large so that in the above set up we have for each $i \in I$ (see (6)),

$$\mathbb{P}(C_H \mid \mathcal{L}_\delta, \mathcal{F}_i) \geq (1 - \varepsilon_1)\mathbf{1}(\mathcal{M}_i).$$  

**Proof.** It follows from the definition of $I$ that on the event $\mathcal{M}_i$, there exists a subset $S$ of $D_i$ with $|S| \geq \varepsilon^2 n$ such that $R_v \leq M$ for each $v \in S$. Using Lemma 2.5, it follows that on $\mathcal{M}_i$,

$$\mathbb{P}(C_H \mid \mathcal{L}_\delta, \mathcal{F}_i) \leq \prod_{v \in S} \mathbb{P}\left(\text{Exp}(0, R_v) \leq R_v - \frac{H}{n}\right) \leq \left(1 - \frac{cH}{n}\right)^{\varepsilon^2 n}$$  

for some constant $c = c(M) > 0$. Clearly, by choosing $H$ sufficiently large (depending on $c$ and $\varepsilon$ and $\varepsilon_1$) one can make the right hand side above smaller than $\varepsilon_1$. This completes the proof of the lemma. □
Remark 2.8. Note that the above proof only relies on the following property of Exponential variables: for each small enough \( \varepsilon \), and for any large enough \( M \), there exists \( H \) such that for all large enough \( n \), \( \mathbb{P}(\text{Exp}(0, R) \in [R - \frac{H}{n}, R]) = \Theta(\frac{1}{\varepsilon n}) \), uniformly for all \( 0 \leq R \leq M \), which is a simple consequence of the fact that the density of Exponential variables is uniformly away from zero on compact sets. Also note that Lemma 2.6 does not rely on any special property of the Exponential distribution.

Note that from Lemma 2.7, it follows that

\[
\mathbb{P}(\mathcal{C}_H \mid \mathcal{L}_\delta) \geq (1 - \varepsilon_1)\mathbb{P}(\mathcal{M}_i).
\]

Thus the proof of Theorem 1.5 is immediate if we could show the existence of \( i \in I \) such that \( \mathcal{M}_i \) holds with probability very close to 1. However we are not quite able to show the latter and instead our basic strategy is to show that for many \( i \)'s in \( I \) the event \( \mathcal{M}_i \) holds with significant probability though not quite close to one. The next few lemmas show why this suffices.

To this end, it is useful to make the following definition. For \( i \in I \), define

\[
\mathcal{B}_i = \left\{ \sum_{v \in D_i} X_v \leq (1 - 2\varepsilon)|D_i| \right\},
\]

where \( \varepsilon \) appears in the statement of Proposition 2.1. The following lemma relates \( \mathcal{B}_i \) to \( \mathcal{M}_i \).

**Lemma 2.9.** There exists a constant \( c > 0 \) such that for each \( i \in I \) we have,

\[
\mathbb{P}(\mathcal{B}_i \cap \mathcal{M}_i^c \mid \mathcal{L}_\delta) \leq e^{-cn}.
\]

**Proof.** Notice that it follows from Lemma 2.5 and Lemma 2.6 that

\[
\mathbb{P}(\mathcal{B}_i \mid \mathcal{L}_\delta, \mathcal{F}_i) \leq e^{-c|D_i|}1(\mathcal{M}_i^c) + 1(\mathcal{M}_i).
\]  

(8)

This implies the lemma as \( |D_i| \) is linear in \( n \) for all \( i \in I \). \( \square \)

Lemma 2.9 tells that we can essentially replace \( \mathcal{M}_i \) by \( \mathcal{B}_i \) in Lemma 2.7, even though \( \mathcal{B}_i \) is not \( \mathcal{F}_i \) measurable. Also observe that by Proposition 2.1 (and definition of \( I \)) we know that it is very likely that at least one of the \( \mathcal{B}_i \)'s hold. In fact we have something stronger.

**Corollary 2.10.** Given \( \delta \), let \( \varepsilon > 0 \) be as in Proposition 2.1. Then,

\[
\mathbb{P}\left( \sum_{i \in I} 1(\mathcal{B}_i) > \varepsilon n \mid \mathcal{L}_\delta \right) \geq 1 - e^{-cn},
\]

for some \( c = c(\delta) > 0 \) and all \( n \) large enough.

**Proof.** On the event \( \{ \sum_{i \in I} 1(\mathcal{B}_i) < \varepsilon n \} \), for \( n \) sufficiently large, we have

\[
\sum_{v \in [0,n]^2} X_v \geq \sum_{i \in I: 1(\mathcal{B}_i) = 0} |D_i| - 2\varepsilon n^2 \geq n^2 - \varepsilon n^2 - 2\varepsilon n^2 > (1 - 4\varepsilon)n^2.
\]

The proof is now complete by Proposition 2.1. \( \square \)
The strategy for completing the proof of Theorem 1.5 now is to check $B_i$'s one by one and control the conditional probability of $C_H$ given a subset of them fails. We need the following lemma ($\delta$ as in Theorem 1.5 will be fixed throughout the sequel).

**Lemma 2.11.** Fix small enough $\varepsilon > 0$. There exists $H = H(\varepsilon)$ sufficiently large such that for all large enough $n$ if $\{i_1, i_2, \ldots, i_k\} \subset I$ is such that for any $1 \leq j \leq k$,

$$\mathbb{P}(B_{i_j} | B_{i_1}^c, \ldots, B_{i_{j-1}}^c, L_\delta) \geq \frac{\varepsilon}{8},$$

then for $1 \leq j \leq k$,

$$\mathbb{P}(C_H | B_{i_j}, B_{i_1}^c, \ldots, B_{i_{j-1}}^c, L_\delta) \geq 1 - \frac{\varepsilon}{2}.$$

Proof of Lemma 2.11 is somewhat technical so we postpone it for the moment. To apply this lemma we need to demonstrate existence of subsets satisfying the hypothesis; and this is the content of the next lemma.

**Lemma 2.12.** Fix small enough $\varepsilon$. Then for all large enough $n$ there exists a subset $J = \{i_1, i_2, \ldots, i_k\} \subset I$ such that we have

(i) $\mathbb{P}(\bigcup_{j \in J} B_j | L_\delta) \geq 1 - \frac{\varepsilon}{2}$.

(ii) For any $1 \leq j \leq k$, we have $\mathbb{P}(B_{i_j} | B_{i_1}^c, \ldots, B_{i_{j-1}}^c, L_\delta) \geq \frac{\varepsilon}{8}$.

We shall postpone this proof too for the moment and complete the proof of Theorem 1.5 first.

**Proof of Theorem 1.5.** Fix $\varepsilon > 0$ and let $H$ be sufficiently large so that the conclusion of Lemma 2.11 holds and let $J \subset I$ be such that the conclusion of Lemma 2.12 holds. Clearly,

$$\mathbb{P}(C_H | L_\delta) \geq \sum_{j=1}^{k} \mathbb{P}(C_H \cap B_{i_j} \cap B_{i_1}^c \cap \cdots \cap B_{i_{j-1}}^c | L_\delta)$$

$$= \sum_{j=1}^{k} \mathbb{P}(C_H | B_{i_j} \cap B_{i_1}^c \cap \cdots \cap B_{i_{j-1}}^c, L_\delta) \mathbb{P}(B_{i_j} \cap B_{i_1}^c \cap \cdots \cap B_{i_{j-1}}^c | L_\delta)$$

$$\geq (1 - \frac{\varepsilon}{2}) \mathbb{P}(\bigcup_i B_i | L_\delta),$$

$$\geq 1 - \varepsilon,$$

where in the last two inequalities we have used Lemmas 2.11 and 2.12. This completes the proof of the theorem. □

It remains to prove Lemmas 2.11 and 2.12. We start with the proof of Lemma 2.12.

**Proof of Lemma 2.12.** We construct the set $J$ in the following manner. Let $J$ denote the collection of all ordered sequences $J = \{j_1, j_2, \ldots, j_\ell\}$ contained in $I$ such that, for any $1 \leq k \leq \ell$, we have

$$\mathbb{P}(B_{j_k} | B_{j_1}^c, \ldots, B_{j_{k-1}}^c, L_\delta) \geq \frac{\varepsilon}{8}.$$
Let \( J_*=\{i_1, i_2, \ldots, i_k\} \) be the sequence in \( J \) that maximizes \( \mathbb{P}(\cup_{j \in J} B_j \mid \mathcal{L}_\delta) \). Suppose now by way of contradiction that \( \mathbb{P}(\cup_{j \in J_*} B_j \mid \mathcal{L}_\delta) < 1 - \frac{\varepsilon}{2} \). The maximality assumption on \( J_* \) implies that for any \( i \notin J_* \) we have
\[
\mathbb{P}(B_i \mid \bigcap_{j \in J_*} B^c_j, \mathcal{L}_\delta) < \frac{\varepsilon}{8}
\]
which implies,
\[
\mathbb{E}\left[\sum_{i \in I} \mathbf{1}(B_i) \mid \bigcap_{j \in J_*} B^c_j, \mathcal{L}_\delta\right] \leq \frac{\varepsilon n}{4},
\]
where the last inequality follows since \(|I| \leq 2n \) [see (6)]. Markov inequality now implies
\[
\mathbb{P}\left[\sum_{i \in I} \mathbf{1}(B_i) \geq \varepsilon n \mid \bigcap_{j \in J_*} B^c_j, \mathcal{L}_\delta\right] \leq \frac{1}{4},
\]
Finally we conclude,
\[
\mathbb{P}\left[\sum_{i \in I} \mathbf{1}(B_i) < \varepsilon n \mid \mathcal{L}_\delta\right] \geq \frac{3}{4} \mathbb{P}(\bigcap_{j \in J_*} B^c_j \mid \mathcal{L}_\delta) > \frac{3\varepsilon}{8},
\]
which contradicts (9). This completes the proof of the lemma. \( \Box \)

Finally we move towards the proof of Lemma 2.11. We first need to set up some more notation. It would be useful to have explicit notation for the sample space and its various projections. Let \( \Omega = [0, \infty)^{[0,n]^2} \) be the product space on which the product measure of exponential vertex weights live. Fix \( i_1, i_2, \ldots, i_k \) such that the hypothesis of Lemma 2.11 holds, i.e., for any \( 1 \leq j \leq k \),
\[
\mathbb{P}(B_{i_j} \mid B^c_{i_1}, \ldots, B^c_{i_{j-1}}, \mathcal{L}_\delta) \geq \frac{\varepsilon}{8}. \tag{10}
\]
For \( \omega \in \Omega \), let \( \omega_j \) denote the projection of \( \omega \) onto the co-ordinates \([0, n]^2 \setminus D_{i_j}, \) i.e., the collection of weights of all the vertices except on the anti-diagonal \( D_{i_j} \). Let \( \Omega_j \) denote the set of all \( \omega_j \)'s. Throughout the sequel for brevity we will naturally identify subsets of \( \Omega_j \) with their pre-image in \( \Omega \) under the projection map. Notice that, whether \( \omega \in B^c_{i_j} \) or not, for \( i \neq i_j \) is a deterministic function of \( \omega_j \). To improve transparency, we shall break the argument into a number of short lemmas.

Let \( \Upsilon_j \subset \Omega_j \) be the set of all \( \omega_j \)'s such that \( \mathbf{1}(B^c_{i_j}(\omega_j)) = 1 \) for \( i = i_1, \ldots, i_{j-1} \). Further, let \( \mathcal{S}_j \subset \Upsilon_j \) be the set of all \( \omega_j \) such that
\[
\mathbb{P}(B_{i_j} \mid \omega_j, \mathcal{L}_\delta) \leq \varepsilon^3. \tag{11}
\]
We now have the following lemma.

**Lemma 2.13.** In the above set-up, there exists \( H \) sufficiently large depending only on \( \varepsilon \), such that for any \( \omega_j \in \Upsilon_j \setminus \mathcal{S}_j \) we have
\[
\mathbb{P}(C_H \cap B_{i_j} \mid \omega_j, \mathcal{L}_\delta) \geq \mathbb{P}(B_{i_j} \mid \omega_j, \mathcal{L}_\delta)(1 - \varepsilon^2). \tag{12}
\]
Proof. For every \( \omega_j \in \Upsilon_j \setminus S_j \) by definition
\[
\mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta) \geq \varepsilon^3.
\]
By (8) this implies \( \mathbf{1}(\mathcal{M}_{ij}(\omega_j)) = 1 \) (notice that \( \mathbf{1}(\mathcal{M}_{ij}) \) is also a deterministic function of \( \omega_j \)). Now by Lemma 2.7 (applied with \( \varepsilon_1 \) replaced by \( \varepsilon^5 \)) it follows that there exists \( H \) sufficiently large such that for any \( \omega_j \in \Upsilon_j \setminus S_j \)
\[
\mathbb{P}(C_H \mid \omega_j, \mathcal{L}_\delta) \geq 1 - \varepsilon^5.
\]
Thus
\[
\mathbb{P}(C_H \cap B_{ij} \mid \omega_j, \mathcal{L}_\delta) \geq \mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta) - \mathbb{P}(C^c_H \mid \omega_j, \mathcal{L}_\delta)
= \mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta)(1 - \frac{\mathbb{P}(C^c_H \mid \omega_j, \mathcal{L}_\delta)}{\mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta)})
\geq \mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta)(1 - \frac{\varepsilon^5}{\varepsilon^3})
= \mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta)(1 - \varepsilon^2).
\]
\( \square \)

Let \( H \) be fixed for the remainder of this section such that the conclusion of Lemma 2.13 holds. We also have the following corollary of Lemma 2.13.

**Corollary 2.14.** In the set-up of Lemma 2.13, we have
\[
\mathbb{P}(C_H \cap B_{ij} \cap (\Upsilon_j \setminus S_j) \mid \mathcal{L}_\delta) \geq (1 - \varepsilon^2)\mathbb{P}(B_{ij} \cap (\Upsilon_j \setminus S_j) \mid \mathcal{L}_\delta).
\]  \( (13) \)

**Proof.** The proof follows by integrating both sides of \( (12) \) over \( \omega_j \in \Upsilon_j \setminus S_j \) (with respect to the conditional density given \( \mathcal{L}_\delta \)). \( \square \)

The final piece of the proof of Lemma 2.11 is provided by the next lemma.

**Lemma 2.15.** In the set-up of Lemma 2.13 we have,
\[
\mathbb{P}(B_{ij} \cap (\Upsilon_j \setminus S_j) \mid \mathcal{L}_\delta) \geq (1 - \varepsilon^2)\mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta).
\]  \( (14) \)

**Proof.** By definition of \( \Upsilon_j \), and \( (10) \), we have
\[
\mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) \geq \frac{\varepsilon}{8}\mathbb{P}(\Upsilon_j \mid \mathcal{L}_\delta).
\]
Using this and \( (11) \) we have
\[
\mathbb{P}(B_{ij} \cap (\Upsilon_j \setminus S_j) \mid \mathcal{L}_\delta) = \mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) - \mathbb{P}(B_{ij} \cap S_j \mid \mathcal{L}_\delta)
= \mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) - \int_{\omega_j \in S_j} \mathbb{P}(B_{ij} \mid \omega_j, \mathcal{L}_\delta) d\mathbb{P}(\omega_j \mid \mathcal{L}_\delta)
\overset{(11)}{=} \mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) - \int_{\omega_j \in S_j} \varepsilon^3 d\mathbb{P}(\omega_j \mid \mathcal{L}_\delta)
\geq \mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) - \varepsilon^2 \mathbb{P}(S_j \mid \mathcal{L}_\delta)
\geq \mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta) - \varepsilon^3 \mathbb{P}(\Upsilon_j \mid \mathcal{L}_\delta)
\geq (1 - 8\varepsilon^2)\mathbb{P}(B_{ij} \cap \Upsilon_j \mid \mathcal{L}_\delta),
\]
completing the proof of the lemma. \( \square \)
Proof of Lemma 2.11 is now immediate.

**Proof of Lemma 2.11.** From (13) and (14) we deduce that
\[
P(C_H \cap B_{ij} \cap \mathcal{Y}_j \mid \mathcal{L}_\delta) \geq (1 - \varepsilon^2)(1 - 8\varepsilon^2)P(B_{ij} \cap \mathcal{Y}_j \mid \mathcal{L}_\delta).
\]
Thus for all small enough \(\varepsilon\) we get,
\[
P(C_H \mid B_{ij}, \mathcal{Y}_j, \mathcal{L}_\delta) \geq 1 - \varepsilon/2,
\]
as required. \(\square\)

3. Anti-Concentration for Restricted Paths in Conditional Environment

Our objective in this section is to prove Theorem 1.6. Recall that \(\delta \in (0, 4)\) is fixed as before. Let us fix \(H > 0\) and \(\varepsilon_A > 0\). Also fix \(A \subseteq [0,n]^2\) which is connected and contains both \(0\) and \(n\) with \(|A| \leq \varepsilon'n^2\). As mentioned before it is convenient to use a standard decoupling property of a collection of i.i.d. exponential variables. Recall the following well-known fact.

**Fact 1.** Let \(X_1, X_2, \ldots, X_m\) be i.i.d. \(\text{Exp}(1)\) variables. Let \(Z = \sum_{i=1}^m X_i\) and \(Y_i = X_i / Z\). Then \(Z \sim \text{Gamma}(m)\) with density proportional to \(e^{-t}t^{m-1}\) on \([0, \infty)\) and is independent of the random vector \((Y_1, \ldots, Y_m)\) having distribution \(F_m\). As a matter of fact, the distribution \(F_m\) is also explicitly known (Dirichlet distribution which is the uniform distribution on the unit positive simplex \(S_m\) of dimension \(m - 1\), i.e.,
\[
S_m = \{(y_1, y_2, \ldots, y_m): y_i \geq 0 \ \forall \ 1 \leq i \leq m, \sum_{i=1}^m y_i = 1\}.
\]
However the latter fact will not be important for us.

The above representation allows us to sample the distribution of \(\{X_v : v \in A\}\) conditional on \(\mathcal{L}_\delta\) in the following way: We think of the field \(\mathbf{X} = (X_v : v \in [0,n]^2) =: (X_{A^c}, X_A)\) as the triple \((X_{A^c}, Y_A, Z_A)\) where \(X_{A^c} := (X_v : v \in A^c)\), and similarly \(X_A := (X_v : v \in A)\), \(Z_A := \sum_{v \in A} X_v\), and \(Y_A := (Y_v : v \in A)\) where \(Y_v = X_v / Z_A\). Thus \(Y_A\) is an element of \(S_{|A|}\) and \(Z_A \sim \text{Gamma}(|A|)\) and by the above fact, all the elements of the triple \((X_{A^c}, Y_A, Z_A)\) are independent of each other.

**Lemma 3.1.** For all \((x_{A^c}, y_A)\) (compatible with \(\mathcal{L}_\delta\), i.e., there exists some \(z_A > 0\), with \((x_{A^c}, y_A, z_A) \in \mathcal{L}_\delta\), we have the following: conditional on \(Y_A = y_A, X_{A^c} = x_{A^c}\) and the event \(\mathcal{L}_\delta\), the random variable \(Z_A\) is distributed as a Gamma\(|A|\) variable conditioned to be in \([0, \theta_{\text{max}}]\) for some constant \(\theta_{\text{max}} = \theta_{\text{max}}(X_{A^c}, Y_A)\). Moreover \(\theta_{\text{max}} \leq \frac{(4 - \delta)n}{L_n(A; Y_A)}\), where \(L_n(A; Y_A)\) denotes the weight of the maximum weight path from \(0\) to \(n\) restricted to lie inside \(A\) and passing through the environment \(Y_A\).

**Proof.** The proof follows by the observation that fixing \(Y_A = y_A, X_{A^c} = x_{A^c}\), \(L_n\) is a non-decreasing continuous function of \(Z_A\) and hence the event \(\mathcal{L}_\delta\) is equivalent to \(Z_A\) being less than some constant \(\theta_{\text{max}}\). Since choosing \(Z_A = \frac{(4 - \delta)n}{L_n(A; Y_A)}\) forces \(L_n \geq L_n(A) = (4 - \delta)n\), the stated upper bound on \(\theta_{\text{max}}\) follows. \(\square\)

We are now ready to prove Theorem 1.6.
Proof of Theorem 1.6. Let \( s = \frac{(4-\delta)n - \frac{H}{n}}{(4-\delta)n} = 1 - \frac{H}{(4-\delta)n} \). Now by Lemma 3.1, it follows that
\[
P \left( L_n(A) \geq (4-\delta)n - \frac{H}{n} \mid \mathcal{L}_\delta, Y_A, X_{A'} \right) \leq \mathbb{E} \left[ P(Z_A \geq s\theta_{\max} \mid \mathcal{L}_\delta, Y_A, X_{A'}) \right]
\]
(15)
where the expectation in the right hand side above is over the distribution of \( Y_A, X_{A'} \) conditional on \( \mathcal{L}_\delta \). To see this notice that \( L_n(A) = L_n(A; Y_A)Z_A \) and hence if \( Z_A < s\theta_{\max} \), then
\[
L_n(A) < L_n(A; Y_A)s\theta_{\max} \leq L_n(A; Y_A)s \frac{(4-\delta)n}{L_n(A; Y_A)} < (4-\delta)n - \frac{H}{n}.
\]
Using Lemma 3.1 and the density of Gamma distribution it follows that
\[
P(Z_A \geq s\theta_{\max} \mid \{Y_v\}) = \int_0^{\theta_{\max}} \frac{e^{-t|A|-1}}{s\theta_{\max}} \, dt \leq \int_0^{\theta_{\max}} e^{-t|A|-1} \, dt.
\]
Doing the change of variable \( t \mapsto \frac{t}{\theta_{\max}} \) we get that the above is upper bounded by
\[
\frac{e^{-s\theta_{\max}} \int_0^1 t^{|A|-1} \, dt}{e^{-s\theta_{\max}} \int_0^s t^{|A|-1} \, dt} \leq \left( 1 - \frac{s}{|A|} \right) \leq \left( e^{\frac{2H\epsilon'}{\delta}} - 1 \right) = O(\epsilon')
\]
where the final inequality follows by taking \( n \) large enough, substituting the value of \( s \) and using the inequality \((1 + x) \leq e^x\). \( \square \)

We end this section with a brief discussion about the stronger notion of delocalization alluded to after the statement of Theorem 1. Note that our proof proceeds by first showing that the polymer weight is concentrated in a \( O(\frac{1}{\delta}) \) window to the left of \( (4-\delta)n \) and then in the current section we showed that any path restricted to lie in a set of size \( o(n^2) \) cannot achieve that level of concentration, implying delocalization. Recalling the definition of \( R_v \) from (7), the first step was shown by establishing the existence of \( \Theta(n) \) many vertices \( v \) with \( R_v \) not too large, on certain anti-diagonals. However to show the stronger notion of delocalization on a fixed anti-diagonal, one has to achieve a finer control and prove a lower bound of the following kind: with high probability there does not exist any small set of vertices on a fixed anti-diagonal with exceptionally small values of \( R_v \) which could cause the polymer to localize on such sets with high probability. We are currently unable to prove such a local anti-concentration statement and achieve the second step above by a more global argument.

We finish with mentioning another related question that we have not investigated in the current paper, namely that of the scaling limit of the whole environment conditional on the lower tail large deviation event. It is perhaps not unnatural to predict the existence, in some sense, of a limiting shape similar to the one which holds typically (see Theorem 1.3), even conditioned on \( \mathcal{L}_\delta \). The results in the current paper suggest that one should expect a family of geodesics in the scaling limit while in the pre-limit, the geodesic which is unique by continuity, should lie close to one such randomly chosen curve, owing to microscopic fluctuations. A program to obtain a shape theorem in such large deviation regimes has been initiated by the authors in [2] in the context of first passage percolation.
4. Extensions to Non-integrable Settings

So far we have been assuming our setting to be the exactly solvable Exponential LPP on \( \mathbb{Z}^2 \). However at this point it is important to note that even though the assumption that the noise variables are exponential made the proof simpler for e.g. the decoupling Fact 1, it was not essential. Our argument, in its core, does not depend on exact solvability and in this section we provide examples of some general settings to which our results can be adapted.

4.1. Last passage percolation on \( \mathbb{Z}^2 \) with general weights.

Consider last passage percolation on \( \mathbb{Z}^2 \) with general i.i.d. vertex weights \( X_v \) coming from some distribution \( F \) on the positive real line. As before let \( L(u, v) \) denote the last passage time from \( u \) to \( v \) for \( u \preceq v \in \mathbb{Z}^2 \). (We shall make use of the same notations as before for other quantities as well). Under some fairly mild moment conditions on \( F \), the law of large numbers result analogous to Theorem 1.3, goes through. The following theorem was proved in [15] (see also [7, 10]).

**Theorem 4.1.** Suppose \( \int_0^\infty (1 - F(x))^{1/2} \, dx < \infty \). Then there exists a function \( G = G_F : \mathbb{R}_+^2 \to \mathbb{R}_+ \) such that for each \( x, y > 0 \) we have

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[L_{\lfloor nx \rfloor, \lfloor ny \rfloor}] = G(x, y).
\]

It is easy to see that such a \( G \) is invariant under swapping of co-ordinates and is super additive, i.e., \( G(x_1, y_1) + G(x_2, y_2) \leq G(x_1 + x_2, y_1 + y_2) \). The last observation implies that the set \( \{(x, y) : G(x, y) = 1\} \) is the graph of a concave function, and it is believed that for a very general class of \( F \), it is strictly concave. This model is also supposed to exhibit KPZ fluctuations of order \( n^{1/3} \), for a general class of \( F \), although it is known only the cases of Exponential, Geometric and Bernoulli distributed weights. Let us restrict ourself to the case where \( F \) is continuous, so that almost surely there is a unique geodesic between any pair of points and as before let \( \Gamma_n \) denote the geodesic between \( 0 \) and \( n \).

One difference from the exactly solvable cases is that it is not rigorously known that \( \Gamma_n \) is concentrated around the diagonal line. An analogue of Theorem 1.2 can however be proved under the assumption of strict concavity of the limit shape and some nice tails of \( F \), and hence is believed to be true for a general class of passage time distributions. Therefore it seems natural to consider the question of localization/ delocalization of geodesics in the lower tail large deviations regime.

As mentioned in Sect. 1.2, the large deviation events are less well understood in the non-integrable setting, with no explicit formulae for the large deviation rate functions unlike the exactly solvable models. However it can be shown that the speed of the lower tail large deviations is of order \( n^2 \) as before. More precisely, let \( F \) satisfy the hypothesis of Theorem 4.1 and let \( \mu = G(1, 1) \). Fix \( \delta \in (0, \mu) \). It can be shown, following the argument outlined in Sect. 1.3 (using exponential concentration below the mean for sums of i.i.d. positive random variables), that

\[
-\infty < \liminf_{n \to \infty} \frac{\log P(L_n \leq (\mu - \delta)n)}{n^2} \leq \limsup_{n \to \infty} \frac{\log P(L_n \leq (\mu - \delta)n)}{n^2} < 0,
\]

(also see [14]).

We shall show, that under certain additional assumptions, the analogue of Theorem 1 remains valid in this setting. Before making a formal statement we shall need to define two classes of probability measures.
**Definition 4.2.** Let \( \mathcal{P} \) denote the class of all probability measures with support \([0, \infty)\), with continuous and positive density that satisfy the hypothesis of Theorem 4.1. Let \( \mathcal{P}_1 \subseteq \mathcal{P} \) denote the class of probability measures with non-increasing density and let \( \mathcal{P}_2 \subseteq \mathcal{P} \) denote the class of all probability measures with log concave density, i.e., density of the form \( \exp(-V(\cdot)) \) where \( V(\cdot) \) is a convex function that is continuously differentiable on \([0, \infty)\) with \( V'(0) > -\infty \).

Our main result in this section is to show that the delocalization result Theorem 1 remains valid in the setting of last passage percolation with general i.i.d. weights as long as the weights come from a distribution in \( \mathcal{P}_1 \) or \( \mathcal{P}_2 \). Recall the notion of an \( \varepsilon \)-cylinder \( \gamma_n^\varepsilon \) around a continuous surjective increasing function \( \gamma : [0, 1] \to [0, 1] \) [see (1)].

**Theorem 3.** Let \( F \) be a probability measure that is either in \( \mathcal{P}_1 \) or in \( \mathcal{P}_2 \). Let \( \mu = G_F(1, 1) \) where \( G_F \) is as in Theorem 4.1. Fix \( \delta \in (0, \mu) \) and \( \varepsilon > 0 \), and set \( L_\delta := \{L_n \leq (\mu - \delta)n\} \). There exists \( \varepsilon' > 0 \) such that for all \( \gamma : [0, 1] \to [0, 1] \) surjective and increasing one has

\[
\mathbb{P}(\Gamma_n \subseteq \gamma_n^\varepsilon' \mid L_\delta) \leq \varepsilon
\]

for all \( n \in \mathbb{N} \).

It will be clear from the proof that the condition that \( F \) is in \( \mathcal{P}_1 \) or in \( \mathcal{P}_2 \) is not optimal, even for our argument. We have not attempted to find the most general class of distributions for which our proof works. With a view of keeping the exposition as simple as possible, our objective was to find a class of distributions that is sufficiently general to be of interest. Observe also that the special case of exponential distribution treated in Theorem 1 falls into both the classes \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \).

The proof of this theorem follows along the same lines as that of Theorem 1. That is, we first show that conditional on \( L_\delta \), the length of the geodesic is concentrated at scale \( 1/n \). As already mentioned in Remark 2.8, this part of the argument does not use any crucial property of the Exponential distribution and will go through for any \( F \) in either \( \mathcal{P}_1 \) or \( \mathcal{P}_2 \). The anti-concentration part however requires more work since we can no longer exploit nice decoupling properties as in Fact 1. We shall prove the following proposition which is Theorem 1.6 restated in this setting. Recall that for \( A \subseteq [0, n]^2 \), \( L_n(A) \) denotes the length of the maximal path between \( \mathbf{0} \) and \( \mathbf{n} \) among all paths completely contained in \( A \).

**Proposition 4.3.** Let \( F \) be a probability distribution in \( \mathcal{P}_1 \) or \( \mathcal{P}_2 \), and consider the set-up in Theorem 3. Let \( H \) and \( \varepsilon_3 > 0 \) be fixed. Then for \( \varepsilon' > 0 \) sufficiently small and for all \( A \subseteq [0, n]^2 \) with \( |A| \leq \varepsilon' n^2 \) we have

\[
\mathbb{P}\left(L_n(A) \geq (\mu - \delta)n - \frac{H}{n} \mid L_\delta\right) \leq \varepsilon_3.
\]

As in the application of Fact 1 in Lemma 3.1, fixing \( n \), we study the field \( X = (X_v : v \in [0, n]^2) = (X_A, X_{A^c}) \) through the triple \((X_{A^c}, Y_A, Z_A)\) where all the elements were defined during the proof of Lemma 3.1.

By hypothesis \((X_v : v \in [0, n]^2)\) are distributed as i.i.d. random variables following a distribution \( F \in \mathcal{P}_1 \cup \mathcal{P}_2 \), say with mean \( m \) and density, \( f(\cdot) \). A simple change of variable shows that conditioning on \((X_{A^c}, Y_A) = (x_{A^c}, y_A)\), the density of \( Z_A \) at any \( z > 0 \) is proportional to

\[
z^{\frac{|A| - 1}{2}} \prod_{v \in A} f(z y_v).
\] (16)
Note that given $X, Y, Z_A$, the quantities $L_n$ and $L_n(A)$, are non-decreasing and strictly increasing functions of $Z_A$ respectively. As before we will call $(x_{A^c}, y_A)$ as compatible with $\mathcal{L}_\delta$, if there exists some $z_A > 0$, such that $(x_{A^c}, y_A, z_A) \in \mathcal{L}_\delta$, (note that this in fact is just a property of $x_{A^c}$.)

Now for any $y_A \in S_{|A|}$ (recall that $S_{|A|}$ is the simplex of dimension $|A| - 1$) and compatible $x_{A^c}$ define

$$\theta_{\text{max}} := \theta_{\text{max}}(y_A, x_{A^c}) := \min \left( \sup \{ \theta : (\theta y_A, x_{A^c}) \in \mathcal{L}_\delta \} , 2m|A| \right),$$

where $\theta y_A = (\theta y_v : v \in A)$ and $m$ is the mean of the distribution $F$. We now have the following lemma analogous to Lemma 3.1.

**Lemma 4.4.** For any $y_A \in S_{|A|}$ and compatible $x_{A^c}$, conditional on $(Y_A = y_A, X_{A^c} = x_{A^c})$ and the events $\{Z_A \leq 2m|A|\}$, and $\mathcal{L}_\delta$: the distribution of $Z_A$ is supported on $[0, \theta_{\text{max}}]$ and has the following density at $z \in (0, \theta_{\text{max}})$:

$$f(z) = \frac{z^{|A|-1} \prod_{i=1}^{|A|} f(y_v)}{\int_{0}^{\theta_{\text{max}}} w^{|A|-1} \prod_{i=1}^{|A|} f(w y_v) dw}.$$

**Proof.** The proof is a straightforward consequence of (16) and the definition of $\theta_{\text{max}}$. \qed

Before the proof of Proposition 4.3 we need another short lemma. Let us abbreviate the event $\{Z_A \leq 2m|A|\}$ by $\mathcal{E}_A$.

**Lemma 4.5.** For any $y_A \in S_{|A|}$ and compatible $x_{A^c}$,

$$\mathbb{P} \left( L_n(A) \geq (\mu - \delta)n - \frac{H}{n} \mid y_A, x_{A^c}, \mathcal{E}_A, \mathcal{L}_\delta \right) \leq \mathbb{P} \left( Z_A \geq \theta_{\text{max}}(1 - \frac{M}{n^2}) \mid y_A, x_{A^c}, \mathcal{E}_A, \mathcal{L}_\delta \right) \quad (18)$$

where $M = \frac{H}{\mu - \delta}$.

**Proof.** Recalling the notation $L_n(A; Y_A)$ from Lemma 3.1, clearly, $L_n(A) = L_n(A; X_A) = Z_A L_n(A; Y_A)$. Thus by definition, if $\mathcal{L}_\delta$ holds,

$$\theta_{\text{max}} L_n(A; Y_A) \leq (\mu - \delta)n. \quad (19)$$

Hence

$$L_n(A; X_A) \geq (\mu - \delta)n - \frac{H}{n} \Rightarrow Z_A L_n(A; Y_A) \geq (\mu - \delta)n - \frac{H}{n},$$

$$\Rightarrow \frac{Z_A}{\theta_{\text{max}}} \geq \frac{(\mu - \delta)n - \frac{H}{n}}{(\mu - \delta)n}.$$ \qed

We are now ready to prove Proposition 4.3.
Proof of Proposition 4.3. Let $s = 1 - \frac{M}{n^2}$ where $M$ is defined in the statement of Lemma 4.5. It follows that

$$\mathbb{P} \left( L_n(A) \geq (\mu - \delta)n - \frac{H}{n} \mid \mathcal{L}_\delta \right)$$

$$\leq \mathbb{P} \left( \mathcal{E}_A \mid \mathcal{L}_\delta \right) \mathbb{P} \left( L_n(A) \geq (\mu - \delta)n - \frac{H}{n} \mid \mathcal{E}_A, \mathcal{L}_\delta \right) + \mathbb{P} \left( \mathcal{E}^c_A \mid \mathcal{L}_\delta \right)$$

$$\leq \mathbb{P} \left( \mathcal{E}_A \mid \mathcal{L}_\delta \right) \mathbb{E} \left[ \mathbb{P} \left( Z_A \geq s\theta_{\max} \mid y_A, \mathbf{x}_{A^c}, \mathcal{E}_A, \mathcal{L}_\delta \right) \right] + \mathbb{P} \left( \mathcal{E}^c_A \mid \mathcal{L}_\delta \right),$$

where the expectation is over the distribution of $(y_A, \mathbf{x}_{A^c})$ conditional on the events $\mathcal{E}_A$ and $\mathcal{L}_\delta$. Using Lemma 4.4, it follows that

$$\mathbb{P} \left( Z_A \geq s\theta_{\max} \mid (y_A, \mathbf{x}_{A^c}), \mathcal{E}_A, \mathcal{L}_\delta \right) = \frac{\int_{s\theta_{\max}}^{\theta_{\max}} z^{|A| - 1} \prod_{v \in A} f(z y_v) \, dz}{\int_{0}^{\theta_{\max}} z^{|A| - 1} \prod_{v \in A} f(z y_v) \, dz}.$$

Doing the change of variable $\frac{z}{\theta_{\max}} \mapsto t$, we get,

$$\frac{\int_{s\theta_{\max}}^{\theta_{\max}} z^{|A| - 1} \prod_{v \in A} f(z y_v) \, dz}{\int_{0}^{\theta_{\max}} z^{|A| - 1} \prod_{v \in A} f(z y_v) \, dz} = \frac{\int_{s}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt}{\int_{0}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt} \leq \frac{\int_{1 - \frac{1}{|A|}}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt}{\int_{1 - \frac{1}{|A|}}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt}.$$  

To complete the proof, we will show that

$$\frac{\int_{s}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt}{\int_{1 - \frac{1}{|A|}}^{1} t^{|A| - 1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt}$$

is small, when $|A| = \epsilon' n^2$ for some small $\epsilon'$. Note that we can ignore the term $t^{|A| - 1}$, in the numerator and denominator since it is $\Theta(1)$ when $t \in [1 - \frac{1}{|A|}, 1]$. Now note that,

$$\int_{1 - \frac{1}{|A|}}^{1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt = \sum_{i=0}^{\frac{n^2}{M|A|} - 1} \int_{1 - (i+1) \frac{M}{n^2}}^{1 - i \frac{M}{n^2}} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt,$$

where to avoid rounding issues we assume $\frac{n^2}{M|A|}$ is an integer. If $F \in \mathcal{P}_1$, then just using monotonicity $i \in \{0, 1, \ldots, \frac{n^2}{M|A|} - 1\}$ we have

$$\int_{1 - (i+1) \frac{M}{n^2}}^{1 - i \frac{M}{n^2}} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt \geq \int_{1 - \frac{M}{n^2}}^{1} \prod_{v \in A} f(t \theta_{\max} y_v) \, dt.$$
and hence the required bound on the RHS in (21) follows. However to prove a similar bound when \( F \in \mathcal{P}_2 \), note that for \( i \in \{0, 1, \ldots, \frac{n^2}{M|A|} - 1 \}, 
\[
\int_{1-(i+1)\frac{M}{n^2}}^{1-i\frac{M}{n^2}} \prod_{v \in A} f(t\theta_{\max,y_v}) dt \geq \left[ \inf_{s \in [1-(i+1)\frac{M}{n^2}, 1-i\frac{M}{n^2}]} \prod_{v \in A} f(s\theta_{\max,y_v}) \right] 
\int_{1-i\frac{M}{n^2}}^{1} \prod_{v \in A} f(t\theta_{\max,y_v}) dt ,
\]
\[
\geq C \int_{1-i\frac{M}{n^2}}^{1} \prod_{v \in A} f(t\theta_{\max,y_v}) dt 
\]
for some \( C > 0 \) (not depending on \( |A| \)) where the last inequality follows from Lemma 4.6 below and that \( \theta_{\max} \leq 2m|A| \). To see this, fixing \( s \in [1-(i+1)\frac{M}{n^2}, 1-i\frac{M}{n^2}] \) we take \( x_A = ((s+i\frac{M}{n^2})\theta_{\max,y_v})_{v \in A} \) and \( t = \frac{s}{s+i\frac{M}{n^2}} \).

Since, by definition, \( \sum_{v \in A} y_v = 1 \), Lemma 4.6 now tells us that
\[
\prod_{v \in A} f(s\theta_{\max,y_v}) \geq e^{-O(m)} ,
\]
which is independent of \( |A| \). Hence whenever \( F \in \mathcal{P}_1 \cup \mathcal{P}_2 \), using the above bound and (22),
\[
\int_{1-\frac{M}{n^2}|A|^{-1}}^{1} \prod_{v \in A} f(t\theta_{\max,y_v}) dt 
\]
\[
= \sum_{i=0}^{\frac{n^2}{M|A|}-1} \int_{1-(i+1)\frac{M}{n^2}}^{1-i\frac{M}{n^2}} \prod_{v \in A} f(t\theta_{\max,y_v}) dt \leq \frac{O(1)}{C \frac{n^2}{M|A|}} = O \left( \frac{|A|}{n^2} \right) ,
\]
where in the last inequality we use the above display and that \( t \in [1-\frac{1}{|A|}, 1] \) implies \( t|A|^{-1} = \Theta(1) \), not depending on \( |A| \). Plugging the above in (20), along with the fact that \( \mathbb{P}(\mathcal{E}_A^c \mid \mathcal{L}_\delta) \) goes to zero completes the proof. To see that \( \mathbb{P}(\mathcal{E}_A^c \mid \mathcal{L}_\delta) \) goes to zero, note that by the FKG inequality
\[
\mathbb{P}(Z_A \geq 2m|A| \mid \mathcal{L}_\delta) \leq \mathbb{P}(Z_A \geq 2m|A|) \]
and the latter goes to zero by law of large numbers as \( \mathbb{E}(Z_A) = m|A| \).

\textbf{Lemma 4.6.} For any density function \( f \) corresponding to a probability measure in \( \mathcal{P}_2 \), there exists \( C > 0 \) such that for any \( x_A \), and any \( 0 < t < 1 \),
\[
\prod_{v \in A} \frac{f(x_v)}{f(t x_v)} < e^{C(1-t)\sum_{v \in A} x_v} .
\]
Proof. Since \( f(\cdot) = -V(\cdot) \), it suffices to show that

\[
\sum_{v \in A} V(tx_v) - \sum_{v \in A} V(x_v) < C(1 - t) \sum_{v \in A} x_v.
\]

Now note that by hypothesis \( \inf_{x \in \mathbb{R}^+} V'(x) = -C > -\infty \). The proof now is a straightforward consequence of mean value theorem. \( \square \)

We can now complete the proof of Theorem 3.

Proof of Theorem 3. Fix a probability distribution \( F \) either in \( \mathcal{P}_1 \) or in \( \mathcal{P}_2 \), \( \delta \in (0, \mu) \) and \( \varepsilon > 0 \). Fix also an increasing surjective function \( \gamma : [0, 1] \to [0, 1] \). Arguing verbatim as in the proof of Theorem 1.5 by Remark 2.8, it follows that there exists \( H \) such that

\[
P\left( L_n \geq (\mu - \delta)n - \frac{H}{n} \mid \mathcal{L}_\delta \right) \geq 1 - \varepsilon/2.
\]

Now by Proposition 4.3, one can choose \( \varepsilon' \) sufficiently small so that \( |\gamma_n^{\varepsilon'}| \leq 2\varepsilon' n^2 \) and

\[
P\left( L_n(\gamma_n^{\varepsilon'}) \geq (\mu - \delta)n - \frac{H}{n} \mid \mathcal{L}_\delta \right) \leq \varepsilon/2.
\]

Combining the above we get \( P(L_n \neq L_n(\gamma_n^{\varepsilon'}) \mid \mathcal{L}_\delta) \geq 1 - \varepsilon. \) \( \square \)

4.2. Last passage percolation on \( \mathbb{Z}^d \). Another setting to which our argument extends in a rather straightforward way is that of directed last passage percolation on higher dimensional Euclidean lattices \( \mathbb{Z}^d \) with i.i.d. weights on the vertices. The last passage percolation model can be defined on \( \mathbb{Z}^d \) for \( d > 2 \), by extending the definition on \( \mathbb{Z}^2 \) in an obvious way. Given positive i.i.d. weights \( \{X_v : v \in \mathbb{Z}^d\} \) one defines the last passage time \( L_n \) from \( 0 = (0, \ldots, 0) \) to \( n = (n, \ldots, n) \) maximizing the weight over all paths that are co-ordinate wise non-decreasing and the weight of a path as before, is the sum of the weights of the vertices on it. Let \( \Gamma_n \) denote the maximizing path (which is unique if \( F \) is continuous which will be the case we shall be restricted to).

The law of large number result (i.e., the analogue of Theorem 4.1) holds in this case provided \( \int_0^\infty (1 - F(x))^{1/d} \, dx < \infty \) (see [15]). Let us define classes of probability distributions \( \mathcal{P}^{(d)}_1 \) and \( \mathcal{P}^{(d)}_2 \) exactly as in Definition 4.2 except that we now also require the above tail condition. Let \( \mu = \mu_F := \lim_{n \to \infty} \frac{\mathbb{E} L_n}{n} \). For any \( \delta \in (0, \mu) \) the following is known about the lower tail large deviation event \( L_n \leq (\mu - \delta)n \)

\[
-\infty < \lim_{n \to \infty} \inf \frac{\log P(L_n \leq (\mu - \delta)n)}{n^d} \leq \lim_{n \to \infty} \sup \frac{\log P(L_n \leq (\mu - \delta)n)}{n^d} < 0.
\]

This can be proved by an easy adaptation of the argument in [13] for first passage percolation. As before let \( \mathcal{L}_\delta := \{L_n \leq (\mu - \delta)n\} \) denote this large deviation event.

Finally for a fixed continuous increasing function \( \gamma : [0, 1] \to [0, 1]^d \) such that \( \gamma(0) = 0 \) and \( \gamma(1) = 1 \), one can define the \( \varepsilon \)-cylinder around \( \gamma \) to be the set of all points at distance at most \( \varepsilon \) from the image of \( \gamma \), i.e., \( \gamma([0, 1]) \) and denote by \( \gamma_n^{\varepsilon} \), the image of the \( \varepsilon \)-cylinder under the scaling map \( x \mapsto nx \). The following is our delocalization result in the higher dimensional setting whose proof is identical to that of Theorem 3 and hence shall be omitted (Note that the strong concentration for \( L_n \) analogous to Theorem 1.5 now occurs at scale \( \frac{1}{n^{d+1}} \)).
Theorem 4. Let $F$ be a probability measure that is either in $P_1$ or in $P_2$. Consider directed last passage percolation on $\mathbb{Z}^d$ ($d > 2$) and let $\mu$ be as above. Fix $\delta \in (0, \mu)$ and $\varepsilon > 0$. There exists $\varepsilon' > 0$ such that for all $\gamma$ as above: one has

$$\mathbb{P}(\Gamma_n \subseteq \gamma_n^{\varepsilon'} | L_\delta) \leq \varepsilon$$

for all $n \in \mathbb{N}$.

Acknowledgements. The authors thank Timo Seppäläinen for pointing out some relevant references, and Ofer Zeitouni for useful conversations. We also thank an anonymous referee for a careful reading of the manuscript as well as several insightful comments. Research of RB is partially supported by an ICTS-Simons Junior Faculty Fellowship and a Ramanujan Fellowship (SB/S2/RJN-097/2017) by Govt. of India. SG’s research was supported by a Miller Research Fellowship. AS is supported by NSF Grant DMS-1352013 and a Simons Investigator grant.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Baik, J., Deift, P., Johansson, K.: On the distribution of the length of the longest increasing subsequence of random permutations. J. Am. Math. Soc 12, 1119–1178 (1999)
2. Basu, R., Ganguly, S., Sly, A.: Upper tail large deviations in first passage percolation. Preprint, arXiv:1712.01255
3. Basu, R., Ganguly, S., Sly, A.: Delocalization of polymers in lower tail large deviation. Preprint, arXiv:1710.11623
4. Basu, R., Sarkar, S., Sly, A.: Coalescence of geodesics in exactly solvable models of last passage percolation. Preprint, arXiv:1704.05219
5. Basu, R., Sidoravicius, V., Sly, A.: Last passage percolation with a defect line and the solution of the slow bond problem. Preprint, arXiv:1408.3464
6. Chow, Y., Zhang, Y.: Large deviations in first-passage percolation. Ann. Appl. Probab. 13, 1601–1614 (2003)
7. Theodore Cox, J., Gandolfi, A., Griffin, P.S., Kesten, H.: Greedy lattice animals I: upper bounds. Ann. Appl. Probab. 3(4), 1151–1169 (1993)
8. Deuschel, J.-D., Zeitouni, O.: Limiting curves for iid records. Ann. Probab. 23, 852–878 (1995)
9. Deuschel, J.-D., Zeitouni, O.: On increasing subsequences of iid samples. Comb. Probab. Comput. 8(03), 247–263 (1999)
10. Gandolfi, A., Kesten, H.: Greedy lattice animals II: linear growth. Ann. Appl. Probab. 4(1), 76–107 (1994)
11. Johansson, K.: Shape fluctuations and random matrices. Commun. Math. Phys. 209(2), 437–476 (2000)
12. Johansson, K.: Transversal fluctuations for increasing subsequences on the plane. Probab. Theory Relat. Fields 116(4), 445–456 (2000)
13. Kesten, H.: École d’Été de Probabilités de Saint Flour XIV-1984, chapter Aspects of first passage percolation, pp. 125–264 (1986)
14. Kesten, H.: First-passage percolation. In: From Classical to Modern Probability, vol. 54 of Progress in Probability, pp. 93–143. Birkhäuser, Basel (2003)
15. Martin, J.B.: Linear growth for greedy lattice animals. Stoch. Process. Their Appl. 98(1), 43–66 (2002)
16. Rost, H.: Nonequilibrium behaviour of a many particle process: density profile and local equilibria. Zeitschrift f. Warsch. Verw. Gebiete 58(1), 41–53 (1981)
17. Seppäläinen, T.: Coupling the totally asymmetric simple exclusion process with a moving interface. Markov Process. Relat. Fields 4(4), 593–628 (1998)
18. Seppäläinen, T.: Large deviations for increasing sequences on the plane. Probab. Theory Relat. Fields 112(2), 221–244 (1998)
19. Talagrand, M.: Concentration of measure and isoperimetric inequalities in product spaces. Publ. Math. l’IHES 81(1), 73–205 (1995)

Communicated by A. Borodin