# ORIENTED FIRST PASSAGE PERCOLATION IN THE MEAN FIELD LIMIT, 2. THE EXTREMAL PROCESS.

NICOLA KISTLER, ADRIEN SCHERTZER, AND MARIUS A. SCHMIDT

ABSTRACT. This is the second, and last paper in which we address the behavior of oriented first passage percolation on the hypercube in the limit of large dimensions. We prove here that the extremal process converges to a Cox process with exponential intensity. This entails, in particular, that the first passage time converges weakly to a random shift of the Gumbel distribution. The random shift, which has an explicit, universal distribution related to modified Bessel functions of the second kind, is the sole manifestation of correlations ensuing from the geometry of Euclidean space in infinite dimensions. The proof combines the multiscale refinement of the second moment method with a conditional version of the Chen-Stein bounds, and a contraction principle.

### 1. INTRODUCTION AND MAIN RESULTS

The model we consider is constructed as follows. We first embed the *n*-dimensional hypercube in  $\mathbb{R}^n$ : for  $e_1, ..., e_n$  the standard basis, we identify the hypercube as the graph  $G_n \equiv (V_n, E_n)$ , where  $V_n = \{0, 1\}^n$  and  $E_n \equiv \{(v, v + e_j) : v, v + e_j \in V, j \leq n\}$ . The set of shortest (directed) paths connecting diametrically opposite vertices, say  $\mathbf{0} \equiv (0, ..., 0)$  and  $\mathbf{1} \equiv (1, ..., 1)$ , is given by

$$\Sigma_n \equiv \{ \pi \in V_{n+1} : \pi_1 = \mathbf{0}, \pi_{n+1} = \mathbf{1}, (\pi_i, \pi_{i+1}) \in E_n, \forall i \le n \}.$$
(1.1)

A graphical rendition is given in Figure 1 below.

2000 Mathematics Subject Classification: 60J80, 60G70, 82B44 Date: August 16, 2018.



FIGURE 1. The 10-dimensional hypercube (left), and two oriented connecting paths (right): blue edges are common to both paths, whereas paths do not overlap on red edges.

Let now  $(\xi_e)_{e \in E}$  be a family of independent standard exponentials, i.e. exponentially distributed random variables with parameter 1, and assign to each oriented path  $\pi \in \Sigma_n$  its *weight* 

$$X_{\pi} \equiv \sum_{k \le n} \xi_{[\pi]_k},$$

where  $[\pi]_i = (\pi_i, \pi_{i+1})$  is the *i*-th edge of the path.

A key question in first passage percolation, FPP for short, concerns the so-called *first* passage time,

$$m_n \equiv \min_{\pi \in \Sigma_n} X_\pi \,, \tag{1.2}$$

namely the smallest weight of connecting paths. The limiting value of  $m_n$  to leading order has been settled by Fill and Pemantle [8], who proved that

$$\lim_{n \to \infty} m_n = 1, \tag{1.3}$$

almost surely.

The "law of large numbers" (1.3) naturally raises questions on fluctuations and weak limits, and calls for a description of the paths with minimal weight. As a first step towards this goal we presented in [11] an alternative, "modern" approach to (1.3) much inspired by the recent advances in the study of Derrida's random energy models (see [9] and references therein) and which relies on the hierarchical approximation to the FPP. In this companion paper we bring the approach to completion by establishing the full limiting picture, i.e. identifying the weak limit of the *extremal process* 

$$\Xi_n \equiv \sum_{\pi \in \Sigma_n} \delta_{n(X_\pi - 1)}$$

**Theorem 1** (Extremal process). Let  $\Xi$  be a Cox process with intensity  $Ze^{x-1}dx$ , where Z is distributed like the product of two independent standard exponentials. Then

$$\lim_{n \to \infty} \Xi_n = \Xi, \tag{1.4}$$

weakly. In particular, it follows for the first passage time  $m_n$  that

$$\lim_{n \to \infty} \mathbb{P}(n(m_n - 1) \le t) = \int_0^\infty \frac{x}{e^{1-t} + x} e^{-x} dx \,.$$
(1.5)

It will become clear in the course of the proof, see in particular Remark 7 below, that the assumption on the distribution of the edge-weights is no restriction: any distribution in the extremality class of the exponentials (i.e. any distribution with similar behavior for small values, to leading order) will lead to the same limiting picture and weak limits. Although not needed, we also point out that the distribution of the mixture is given by  $f(z) = 2z^2 K_0(2\sqrt{z})$ , with  $K_0$  a modified Bessel function of the second kind.

What lies behind the onset of the Cox processes is a *decoupling* whose origin can be traced back to the high-dimensional nature of the problem at hand. Indeed, the following mechanism, depicted in Figure 2 below, holds with overwhelming probability in the limit  $n \to \infty$  first, and  $r \to \infty$  next: Walkers connecting **0** to **1** through paths of minimal weight may share at most the first r steps of their journey. Yet, and crucially: whenever they depart from one another ('branch off'), they cannot meet again until they lie at distance at most r from the target. If meeting happens, they must continue on the same path (no further branching is possible). The long stretches during which optimal paths do not overlap are eventually responsible for the Poissonian component of the extremal process, whereas the mixing is due to the relatively short stretches of tree-like (early and late) evolution of which the system keeps persistent memory. The picture is thus very reminiscent of the extremes of branching Brownian motion [BBM], see [2] and references therein. More specifically, the extremal process of FPP on the hypercube can be (partly) seen as the "gluing together" of two extremal processes of BBM in the weak correlation regime as studied by Bovier and Hartung [3, 4], see also [5, 6, 7].

Acknowledgements. It is our pleasure to thank Ralph Neininger for much needed guidance in the field of contraction methods and distributional fixed points.

### 2. Strategy of proof

The approach amounts to exploiting the insights on the physical mechanisms summarized in Figure 2. Specifically, we will check convergence of intensity and avoidance functions of the extremal process. To see how this comes about, we lighten notation by setting, for  $A \subset \mathbb{R}$  a generic subset and  $\pi$  an oriented path,

$$I_{\pi}(A) \equiv \delta_{n(X_{\pi}-1)}(A)$$
, and  $\Xi_n(A) \equiv \sum_{\pi \in \Sigma_n} I_{\pi}(A)$ .

We then claim that with Z as in Theorem 1, and A a finite union of bounded intervals:

• Convergence of the intensity:

$$\lim \mathbb{E} \Xi_n(A) \xrightarrow[n \to \infty]{} \mathbb{E} \int_A Z e^{x-1} dx = \int_A e^{x-1} dx \,.$$
(2.1)



FIGURE 2. Four extremal paths. Remark in particular the tree-like evolution close to 0 and 1 (blue edges) and the (comparatively) longer stretch where paths share no common edge (red). This should be contrasted with the low-dimensional scenario: "loops" in the core of the hypercube, as depicted in Figure 1, become less and less likely as the dimension grows.

• Convergence of the avoidance function:

$$\mathbb{P}\left(\Xi_n(A)=0\right) \xrightarrow[n\to\infty]{} \mathbb{P}\left(\Xi(A)=0\right) = \mathbb{E}\left[\exp\left(-Z\int_A e^{x-1}dx\right)\right].$$
(2.2)

Theorem 1 then immediately follows in virtue of Kallenberg's Theorem [10, Theorem 4.15]. The proof of the claim on the intensity is rather straightforward: it only requires tail-estimates which we now state for they will be constantly used throughout the paper. (The simple proof may be found in [11, Lemma 5]).

**Lemma 2.** Let  $\{\xi_i\}_{i\leq n}$  be independent standard exponentials, and set  $X_n \equiv \sum_{i=1}^n \xi_i$ . Then

$$\mathbb{P}(X_n \le x) = (1 + K(x, n)) \frac{e^{-x} x^n}{n!},$$
(2.3)

for x > 0 and with the error-term satisfying  $0 \le K(x, n) \le e^x x/(n+1)$ .

Armed with these estimates we can proceed to the short proof of (2.1). Here and below, we will always consider sets of the form  $A = (-\infty, a]$ ,  $a \in \mathbb{R}$ . This is enough for our purposes since the general case follows by additivity. It holds:

$$\begin{split} \mathbb{E}\Xi_{n}(A) &= \sum_{\pi \in \Sigma_{n}} \mathbb{P}\left(n(X_{\pi} - 1) \leq a\right) \\ &= n! \mathbb{P}\left(n(X_{\pi^{*}} - 1) \leq a\right) \qquad (\text{symmetry}, \ \pi^{*} \in \Sigma_{n} \text{ is arbitrary}) \\ &= n! \left\{1 + K\left(1 + \frac{a}{n}, n\right)\right\} e^{-1 - \frac{a}{n}} \left(\left(1 + \frac{a}{n}\right)^{+}\right)^{n} (n!)^{-1} \qquad (\text{Lemma 2}) \qquad (2.4) \\ &= (1 + o_{n}(1))e^{-1 + a} \\ &= (1 + o_{n}(1)) \int_{A} e^{x - 1} dx, \end{split}$$

as claimed. Convergence of the intensity (2.1) is thus already settled.

Contrary to convergence of the intensity, convergence of avoidance functions (2.2) will require a fair amount of work. This will be split in a number of intermediate steps. The main ingredient is a conditional version of the Chen-Stein bounds:

**Theorem 3** (Conditional Chen-Stein Method). Consider a probability space  $(\Omega, \mathscr{F}, \mathbb{P})$ , a sigma-algebra  $\mathcal{F} \subset \mathscr{F}$ , a finite set I, and a family  $(X_i)_{i \in I}$  of Bernoulli random variables issued on this space. Let furthermore

$$W = \sum_{i \in I} X_i$$
 and  $\lambda = \sum_{i \in I} \mathbb{E}(X_i | \mathcal{F}).$ 

Finally, consider a random variable  $\widehat{W}$  with the property that its law conditionally upon  $\mathcal{F}$  is Poisson, i.e.  $\mathcal{L}(\widehat{W}|\mathcal{F}) = \text{Poi}(\lambda)$ . It then holds:

$$d_{TV|\mathcal{F}}(W,\widehat{W}) \le \sum_{i \in I} \mathbb{E}(X_i|\mathcal{F})^2 + \sum_{i \in I} \sum_{j \in N_i} (\mathbb{E}(X_i|\mathcal{F})\mathbb{E}(X_j|\mathcal{F}) + \mathbb{E}(X_iX_j|\mathcal{F})), \quad (2.5)$$

where

$$d_{TV|\mathcal{F}}(W,\widehat{W}) \equiv \sup_{A \in \mathscr{F}} \left( \mathbb{P}_W(A|\mathcal{F}) - \mathbb{P}_{\widehat{W}}(A|\mathcal{F}) \right)$$

is the total variation distance conditionally upon  $\mathcal{F}$ . Finally,  $N_i, i \in I$  is a collection of conditionally dissociating neighborhoods, i.e. with the property that  $X_i$  and  $\{X_j : j \in (N_i \cup \{i\})^c\}$  are independent, conditionally upon  $\mathcal{F}$ .

Theorem 3 is a variant of the classical Chen-Stein method which is tailor-suited to our purposes. Since we haven't found in the literature any similar statement, we provide the rather short proof in the appendix for completeness.

In order to prove convergence of the avoidance functions, we will apply Theorem 3 by conditioning on the left- and rightmost regions of Figure 2, namely those regions where tree-like evolutions eventually kick in. Specifically, in order to apply the conditional Chen-Stein, we make the following choices:

- a)  $I \equiv \Sigma_n$ , the set of admissible (oriented) paths connecting **0** to **1**.
- b)  $\mathcal{F}$  is the sigma-algebra generated by the weights of edges at distance at most r from **0** or **1**, to wit

$$\mathcal{F} = \mathcal{F}_{r,n} \equiv \sigma(\xi_e : e = (u, v) \in E, \min\{d(u, \mathbf{0}), d(v, \mathbf{0})\} \in [0, r) \cup [n - r, n))$$

- c) The family of Bernoulli r.v.'s is given by  $(I_{\pi}(A))_{\pi \in \Sigma_n}$ .
- d) The (random) Poisson-parameter is

$$\lambda = \lambda_{r,n}(A) \equiv \sum_{\pi \in \Sigma_n} \mathbb{E}\left[I_{\pi}(A) \mid \mathcal{F}_{r,n}\right]$$

e) The dissociating neighborhoods are given, for  $\pi \in \Sigma_n$ , by

$$N_{\pi} \equiv \{\pi' \in \Sigma_n \setminus \{\pi\} : \exists i \in \{r+1, .., n-r\} \text{ s.t. } [\pi]_i = [\pi']_i\}$$

A first, fundamental observation concerns item d), namely the weak convergence of the Poisson-parameter in the double limit  $n \to \infty$  first and  $r \to \infty$  next. This is an instructive warm-up computation which we now explain.

Denote the set of all pairs of paths leading r-steps away from the start/end respectively, and which can be part of an oriented path from **0** to **1** by

$$\mathcal{V}_{r,n} = \{ (x,y) \in V^{r+1} \times V^{r+1} : x_1 = \mathbf{0}, d(x_{r+1},\mathbf{0}) = r, d(y_1,\mathbf{1}) = r, y_{r+1} = \mathbf{1}, \\ y_1 - x_{r+1} \in V, (x_i, x_{i+1}), (y_i, y_{i+1}) \in E, \forall i \le r \}.$$
(2.6)

Note that  $y_1 - x_{r+1} \in V$  is equivalent to there being a directed path from **0** to **1** containing x and y. For  $(x, y) \in \mathcal{V}_{r,n}$  we define the set of paths connecting x and y by

$$\Sigma_{x,y} \equiv \{ \pi' \in V^{n-2r+1} : \exists \pi \in \Sigma_n \text{ s.t. } ([\pi]_i)_{i \le r} = ([x]_i)_{i \le r} \text{ and} \\ ([\pi]_i)_{r < i \le n-r} = ([\pi']_i)_{r < i \le n-r}, ([\pi]_i)_{i > n-r} = ([y]_i)_{i > n-r} \}.$$

$$(2.7)$$

By definition,

$$\lambda_{r,n}(A) = \sum_{\pi \in \Sigma_n} \mathbb{P}\left(n(X_{\pi} - 1) \le a \middle| \mathcal{F}_{r,n}\right) = \sum_{(x,y) \in \mathcal{V}_{r,n}} \sum_{\pi' \in \Sigma_{x,y}} \mathbb{P}\left(\sum_{i=1}^{n-2r} \xi_{[\pi']_i} \le 1 + \frac{a}{n} - \sum_{i=1}^r \xi_{[x]_i} + \xi_{[y]_i} \middle| \mathcal{F}_{r,n}\right).$$
(2.8)

Shorten

$$X_{x,y} \equiv \sum_{i=1}^{r} \xi_{[x]_i} + \xi_{[y]_i}$$

By Lemma 2, and since  $|\Sigma_{x,y}| = (n-2r)!$ , the r.h.s. of (2.8) equals

$$\sum_{(x,y)\in\mathcal{V}_{r,n}} \left(1 + K(1 + \frac{a}{n} - X_{x,y}, n - 2r)\right) \exp\left(-1 - \frac{a}{n} + X_{x,y}\right) \left(\left(1 + \frac{a}{n} - X_{x,y}\right)^+\right)^{n-2r}.$$
(2.9)

By the tail-estimates from Lemma 2, the following holds

$$K\left(1 + \frac{a}{n} - X_{x,y}, n - 2r\right) \le \frac{2e^2}{n - 2r},$$

for all non-zero summands, and  $n \ge a$ . Remark that there are  $O(n^{2r})$  such summands, while r and a are fixed: one easily checks that dropping all summands where  $X_{x,y} >$ 

 $(\ln n)^2/n$  only causes a deterministically vanishing error, hence

$$(2.9) = (1 + o_n(1)) \left( o_n(1) + e^{-1} \sum_{(x,y) \in \mathcal{V}_{r,n}} \mathbb{1}_{\{X_{x,y} \le \frac{(\ln n)^2}{n}\}} \exp\left((n - 2r) \ln\left(1 + \frac{a}{n} - X_{x,y}\right)^+\right) \right)$$
$$= (1 + o_n(1)) \left( o_n(1) + e^{-1+a} \sum_{(x,y) \in \mathcal{V}_{r,n}} \exp\left(-nX_{x,y}\right) \right)$$
$$= (1 + o_n(1)) \left( o_n(1) + e^{-1+a} \sum_{(x,y) \in \mathcal{V}_{r,n}} \exp\left(-n\sum_{i=1}^r \left(\xi_{[x]_i} + \xi_{[y]_i}\right)\right) \right).$$
(2.10)

the second step by Taylor-expanding the logarithm around 1 to first order, and the third by definition.

We now address the sum on the r.h.s. of (2.10), on which we perform the aforementioned double limit  $n \to \infty$  first and  $r \to \infty$  next. The upshot is summarized in Proposition 4 below, whose proof — via a contraction argument — is deferred to Section 3.1. To formulate, we need some additional notation: for  $\pi_1, ..., \pi_{i-1} \in \mathbb{N}$  and  $i \leq r$  we denote by

$$(\eta_{\pi_1,\dots,\pi_{i-1},\pi_i})_{\pi_i\in\mathbb{N}}, \text{ and } (\tilde{\eta}_{\pi_1,\dots,\pi_{i-1},\pi_i})_{\pi_i\in\mathbb{N}}$$

independent Poisson point processes [PPP] with intensity  $\mathbb{1}_{\mathbb{R}^+} dx$ , and set

$$Z_r \equiv \sum_{\pi \in \mathbb{N}^r} \exp\left(-\sum_{j=1}^r \eta_{\pi_1 \pi_2 \dots \pi_j}\right), \quad \widetilde{Z}_r \equiv \sum_{\pi \in \mathbb{N}^r} \exp\left(-\sum_{j=1}^r \widetilde{\eta}_{\pi_1 \pi_2 \dots \pi_j}\right).$$
(2.11)

**Proposition 4.** (The double weak-limit).

• n-convergence: the following weak limit, to fixed r, holds:

$$\lim_{n \to \infty} \sum_{(x,y) \in \mathcal{V}_{r,n}} \exp -n \sum_{l=1}^{r} \left( \xi_{[x]_l} + \xi_{[y]_l} \right) = Z_r \times \widetilde{Z}_r \,.$$

• r-convergence:  $Z_r$  and  $\widetilde{Z}_r$  weakly converge, as  $r \to \infty$ , to independent standard exponentials.

The n-convergence is a key ingredient in Figure 2 above. Indeed, remark that both limits  $Z_r$  and  $\tilde{Z}_r$  are constructed outgoing from *hierarchical*<sup>1</sup> superpositions of PPP: this accounts for the somewhat surprising fact that close to **0** and **1** only tree-like structures contribute to the extremal process in the mean field limit.

Proposition 4 and (2.10) steadily imply convegence of the Poisson-parameter:

<sup>&</sup>lt;sup>1</sup>Superpositions of PPP such as those involved in (2.11) are ubiquitous in the Parisi theory of mean field sping glasses, see [9] and references, where they are referred to as *Derrida-Ruelle cascades*. Although no knowledge of the Parisi theory is assumed/needed, our approach to the oriented FPP in the limit of large dimensions heavily draws on ideas which have recently crystallised in that field.

**Corollary 5.** With the above notations,

$$\lim_{r \to \infty} \lim_{n \to \infty} \lambda_{r,n}(A) = Z \int_A e^{x-1} dx,$$

weakly.

We now come back to the main task of proving (2.2), convergence of the avoidance functions. The line of reasoning goes as follows: recalling that  $\Xi_n(A) = \sum_{\pi \in \Sigma_n} I_{\pi}(A)$ , we write

$$|\mathbb{P}(\Xi_n(A) = 0) - \mathbb{P}(\Xi(A) = 0)|$$

$$= |\mathbb{E}\mathbb{P}(\Xi_n(A) = 0 | \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A) = 0 | Z)|$$

$$\leq |\mathbb{E}\mathbb{P}(\Xi_n(A) = 0 | \mathcal{F}_{r,n}) - \mathbb{P}(\operatorname{Poi}(\lambda_{r,n}(A)) = 0 | \mathcal{F}_{r,n})|$$

$$+ |\mathbb{E}\mathbb{P}(\operatorname{Poi}(\lambda_{r,n}(A)) = 0 | \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A) = 0 | Z)|,$$

$$(2.12)$$

by the triangle inequality. By convexity, one has

$$|\mathbb{EP} (\Xi_n(A) = 0 | \mathcal{F}_{r,n}) - \mathbb{P} (\operatorname{Poi} (\lambda_{r,n}(A)) = 0 | \mathcal{F}_{r,n})| \leq \mathbb{E} |\mathbb{P} (\Xi_n(A) = 0 | \mathcal{F}_{r,n}) - \mathbb{P} (\operatorname{Poi} (\lambda_{r,n}(A)) = 0 | \mathcal{F}_{r,n})| \leq \mathbb{E} d_{TV,\mathcal{F}_{r,n}} (\Xi_n(A), \operatorname{Poi} (\lambda_{r,n}(A))) =: \operatorname{CS}(r, n), \quad \text{say.}$$

$$(2.13)$$

Furthemore, by definition

$$\begin{aligned} \left| \mathbb{EP} \left( \operatorname{Poi} \left( \lambda_{r,n}(A) \right) = 0 \mid \mathcal{F}_{r,n} \right) - \mathbb{EP} \left( \Xi(A) = 0 \mid Z \right) \right| \\ &= \left| \mathbb{E} \left( e^{-\lambda_{r,n}(A)} - e^{-Z \int_A e^{x-1} dx} \right) \right| \\ &=: \operatorname{P}(r,n), \quad \text{say.} \end{aligned}$$
(2.14)

It thus follows from (2.12), (2.13) and (2.14) that

$$\left|\mathbb{P}\left(\Xi_n(A)=0\right) - \mathbb{P}\left(\Xi(A)=0\right)\right| \le \operatorname{CS}(r,n) + \operatorname{P}(r,n).$$
(2.15)

The second term is easily seen to vanish thanks to the convergence of the Poissonparameter: it follows from Corollary 5 and weak limit that

$$\lim_{r \to \infty} \lim_{n \to \infty} \mathcal{P}(r, n) = 0.$$
(2.16)

We finally claim that the first term in (2.15), the "Chen-Stein term", also vanishes in the considered double-limit, to wit:

$$\lim_{r \to \infty} \lim_{n \to \infty} \operatorname{CS}(r, n) = 0.$$
(2.17)

This claim is proved in Section 3.2 as an application of the conditional Chen-Stein method.

Combining (2.16) and (2.17) we thus obtain convergence of the avoidance function: since this was the last missing ingredient, our main Theorem 1 follows.

## 3. Proofs

3.1. The double weak-limit. The goal of this section is to prove Proposition 4. We first address the n-convergence, which states that

$$\lim_{n \to \infty} \sum_{(x,y) \in \mathcal{V}_{r,n}} \exp\left(-n \sum_{l=1}^r \xi_{[x]_l} + \xi_{[y]_l}\right) = Z_r \times \widetilde{Z}_r,\tag{3.1}$$

weakly, where  $Z_r, \tilde{Z}_r$  are defined in (2.11). The idea here is to enlarge the set of paths over which the sum is taken, as this enables a useful decoupling. Precisely, consider the set of directed paths of length r from **0**,

$$\mathcal{V}_{r,n}^{\leftarrow} = \{ x \in V^{r+1} : x_1 = \mathbf{0}, d(x_{r+1}, \mathbf{0}) = r, [x]_i \in E, \forall i \le r \},$$
(3.2)

and respectively to 1:

$$\mathcal{V}_{r,n}^{\to} = \{ y \in V^{r+1} : y_{r+1} = \mathbf{1}, d(y_1, \mathbf{1}) = r, [y]_i \in E, \forall i \le r \}.$$
(3.3)

One easily checks that

$$\left|\mathcal{V}_{r,n}^{\leftarrow} \times \mathcal{V}_{r,n}^{\rightarrow} \setminus \mathcal{V}_{r,n}\right| = O(n^{2r-1}).$$
(3.4)

We split the sum over the larger subset into a sum over  $\mathcal{V}_{r,n}$  and a "rest-term":

$$\sum_{(x,y)\in\mathcal{V}_{r,n}^{\rightarrow}\times\mathcal{V}_{r,n}^{\leftarrow}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}+\xi_{[y]_{l}}\right) =$$

$$=\sum_{(x,y)\in\mathcal{V}_{r,n}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}+\xi_{[y]_{l}}\right)$$

$$+\sum_{(x,y)\in(\mathcal{V}_{r,n}^{\rightarrow}\times\mathcal{V}_{r,n}^{\leftarrow})\setminus\mathcal{V}_{r,n}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}+\xi_{[y]_{l}}\right).$$
(3.5)

and claim that the term on the r.h.s. vanishes in probability. Indeed, by a simple computation involving the moment generating function of the exponential distribution,

$$\mathbb{E}\left|\sum_{(x,y)\in(\mathcal{V}_{r,n}^{\to}\times\mathcal{V}_{r,n}^{\leftarrow})\setminus\mathcal{V}_{r,n}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}+\xi_{[y]_{l}}\right)\right| = \left|\left(\mathcal{V}_{r,n}^{\to}\times\mathcal{V}_{r,n}^{\leftarrow}\right)\setminus\mathcal{V}_{r,n}\right|(n+1)^{-2r}$$

$$\xrightarrow{(3.4)}{0}, \quad n\to\infty.$$
(3.6)

It thus follows from Markov's inequality that the contribution of paths in  $(\mathcal{V}_{r,n}^{\rightarrow} \times \mathcal{V}_{r,n}^{\leftarrow}) \setminus \mathcal{V}_{r,n}$  is irrelevant for our purposes: the weak limit when summing over  $\mathcal{V}_{r,n}$ , and that when summing over  $\mathcal{V}_{r,n}^{\rightarrow} \times \mathcal{V}_{r,n}^{\leftarrow}$  coincide, provided one of them exists. On the other hand, the sum over the enlarged set of paths "decouples" into two independent identically distributed terms:

$$\sum_{(x,y)\in\mathcal{V}_{r,n}^{\to}\times\mathcal{V}_{r,n}^{\leftarrow}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}} + \xi_{[y]_{l}}\right) = \sum_{x\in\mathcal{V}_{r,n}^{\to}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}\right) \sum_{y\in\mathcal{V}_{r,n}^{\leftarrow}} \exp\left(-n\sum_{l=1}^{r}\xi_{[y]_{l}}\right).$$
(3.7)

The n-convergence will therefore follow as soon as we show that

$$Z_{r,n} \equiv \sum_{x \in \mathcal{V}_{r,n}^{\to}} \exp\left(-n\sum_{l=1}^{r} \xi_{[x]_l}\right) \xrightarrow[n \to \infty]{} \sum_{\pi \in \mathbb{N}^r} \exp\left(\sum_{l=1}^{r} -\eta_{\pi_1 \pi_2 \dots \pi_j}\right) \equiv Z_r \qquad (3.8)$$

holds weakly. This will be done by induction on r. The base-case r = 1 is addressed in

**Lemma 6.** Consider  $\eta \equiv \sum_{i \in \mathbb{N}} \delta_{\eta_i}$  a PPP  $(\mathbb{1}_{\mathbb{R}^+} dx)$  and independent standard exponentials  $(\xi_i)_{i \in \mathbb{N}}$ . It then holds:

$$\sum_{i=1}^{n} \delta_{\xi_i n} \underset{n \to \infty}{\longrightarrow} \eta \tag{3.9}$$

weakly. Furthermore, the following weak limit holds:

$$\sum_{i=1}^{n} \exp\left(-\xi_{i}n\right) \xrightarrow[n \to \infty]{} \sum_{i \in \mathbb{N}} \exp\left(-\eta_{i}\right) .$$
(3.10)

**Remark 7.** In virtue of Lemma 6, Theorem 1 holds for any choice of edge-weights falling in the universality class of the exponential distribution, i.e. for which (3.9) holds.

Proof of Lemma 6. Claim (3.9) is a classical result in extreme value theory. We omit the elementary proof. As for the second claim: it is steadily checked (e.g. by Markov's inequality) that the sum on the l.h.s. of (3.10) is almost surely finite. In order to prove (3.10) it thus suffices to compute the Laplace transform of the two sums. For  $t \in \mathbb{R}^+$ , since the  $\xi$ 's are independent, we have:

$$\mathbb{E} \exp -t \sum_{i=1}^{n} e^{-\xi_{i}n} = \mathbb{E} \left( e^{te^{-\xi_{1}n}} \right)^{n} = \left( 1 + \int_{0}^{+\infty} e^{-x} (e^{te^{-xn}} - 1) dx \right)^{n} = \left( 1 + \frac{1}{n} \int_{0}^{+\infty} e^{-u/n} (e^{te^{-u}} - 1) du \right)^{n},$$
(3.11)

the second equality by change of variable. But  $e^{-u/n}(e^{te^{-u}}-1) \leq (e^{te^{-u}}-1)$ , which is integrable, hence by dominated convergence we have that the r.h.s. of (3.11) converges, as  $n \uparrow \infty$ , to the limit

$$\exp\left(\int_{0}^{+\infty} (e^{-te^{-x}} - 1)dx\right) = \mathbb{E}\exp-t\sum_{i\in\mathbb{N}}e^{-\eta_i},\tag{3.12}$$

where the last equality follows by a simple computation: (3.10) is therefore settled.

For the n-convergence, we will work with the Prohorov metric, which we recall is defined as follows: for  $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$  two probability measures, the Prohorov distance is given by

$$d_{\mathcal{P}}(\mu,\nu) \equiv \inf \left\{ \epsilon > 0 : \mu(A) \le \nu(A^{\epsilon}) + \epsilon, \ \forall A \subset \mathbb{R} \text{ closed} \right\}$$

where  $A^{\epsilon} \equiv \{x \in \mathbb{R} : d(A, x) \leq \epsilon\}$  is the  $\epsilon$ -neighborhood of the set A. It is a classical fact that the Prohorov distance metricizes weak convergence. We also recall the following

implication, as it will be used at different occurences: for two r.v.s X, Y, slightly abusing notation, one has:

$$\mathbb{P}(|X - Y| > \varepsilon) \le \varepsilon \Rightarrow d_{\mathbb{P}}(X, Y) \le \varepsilon.$$
(3.13)

In fact,  $\mathbb{P}(|X - Y| > \varepsilon) \le \varepsilon$  implies that for  $A \subset \mathbb{R}$ ,

$$\mathbb{P}(X \in A) \le \mathbb{P}(X \in A, |X - Y| \le \varepsilon) + \mathbb{P}(|X - Y| > \varepsilon) \le \mathbb{P}(Y \in A^{\varepsilon}) + \varepsilon$$
(3.14)

from which  $d_P(X, Y) \leq \varepsilon$  follows, settling (3.13).

We now proceed to the induction step: we thus assume that  $Z_{r,n}$  converges weakly to  $Z_r$  for some  $r \in \mathbb{N}$  and show how to deduce that  $Z_{r+1,n}$  converges weakly to  $Z_{r+1}$ . First, we observe that by definition

$$Z_{r+1,n} = \sum_{i \le n} \exp\left(-n\xi_{(\mathbf{0},e_i)}\right) \sum_{x \in \mathcal{V}_{r+1,n}^{\rightarrow}: x_2 = e_i} \exp\left(-n\sum_{l=2}^{r+1} \xi_{[x]_l}\right)$$
  
$$= \sum_{i \le n} \exp\left(-n\xi_{(\mathbf{0},e_i)}\right) \times Z_{r,n}^{e_i},$$
(3.15)

changing notation for the second sum to lighten exposition.

We claim that it suffices to consider small  $\xi$ -values in the first sum. Precisely, let  $\varepsilon > 0$ , set  $K_{\varepsilon} = -2 \ln \varepsilon$ , and restrict the first sum to those  $\xi's$  such that  $\xi_{(0,e_i)} \leq K_{\varepsilon}/n$ . We claim that this causes only an  $\varepsilon$ -error in Prohorov distance, to wit

$$\sup_{n,r} d_{\mathcal{P}}\left(Z_{r+1,n}, \sum_{i \le n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)} \le K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i}\right) \le \varepsilon.$$
(3.16)

In fact, for the contribution of large  $\xi's$ , it holds:

$$\mathbb{P}\left(\sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_{i})}>K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_{i})}} \times Z_{r,n}^{e_{i}} > \varepsilon\right) \\
\leq \frac{1}{\varepsilon} \mathbb{E}\left[\sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_{i})}>K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_{i})}} \times Z_{r,n}^{e_{i}}\right] \\
= \frac{n}{\varepsilon} \mathbb{E}\left[\mathbb{1}_{\{\xi_{(\mathbf{0},e_{i})}>K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_{i})}}\right] \times \mathbb{E}\left[Z_{r,n}^{e_{i}}\right],$$
(3.17)

the first step by Markov inequality, and the second by independence. Computing explicitly the above expectations yields that the r.h.s. of (3.17) is at most

$$\frac{n}{\varepsilon} \int_{K_{\varepsilon}/n}^{\infty} e^{-(n+1)x} dx \times \frac{(n-1)!}{(n-r-1)!} (n+1)^{-r} \le \frac{\exp -K_{\varepsilon}}{\varepsilon} = \varepsilon,$$
(3.18)

since  $K_{\varepsilon} = -2 \ln \varepsilon$ . This settles (3.16).

Consider now the permutation p of  $\{1, ..., n\}$  such that  $(\xi_{p(i)})_{i \leq n}$  is increasing, and set  $\hat{K}_{\varepsilon} \equiv \lceil K_{\varepsilon}/\varepsilon \rceil$ . We clearly have

$$Z_{n,r+1} \ge \sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}} .$$

$$(3.19)$$

for  $\geq \hat{K}_{\varepsilon}$ . On the other hand,

$$\mathbb{P}\left(Z_{n,r+1} \ge \sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}} + \varepsilon\right) \\
\le \mathbb{P}\left(Z_{n,r+1} \ge \varepsilon + \sum_{i \le n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)} \le K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i}\right) \\
+ \mathbb{P}\left(\sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}} \le \sum_{i \le n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)} \le K_{\varepsilon}/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i}\right) \tag{3.20}$$

While the first term is at most  $\varepsilon$  by (3.17) and (3.18), the second term equals

$$\mathbb{P}\left(\#\{i \le n : \xi_{(\mathbf{0},e_i)} \le K_{\varepsilon}/n\} > \hat{K}_{\varepsilon}\right) \\
\le n \mathbb{P}(\xi_{(\mathbf{0},e_1)} \le K_{\varepsilon}/n) / \hat{K}_{\varepsilon} \le K_{\varepsilon} / \hat{K}_{\varepsilon} \le \varepsilon,$$
(3.21)

the first estimate by Markov inequality and the second using  $(1 - e^{-x}) \leq x$ .

All in all, in virtue of (3.13), the above considerations imply that

$$\sup_{n,r} d_{\mathcal{P}} \left( Z_{n,r+1}, \sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p}(i)} \right) \le 2\varepsilon .$$
(3.22)

A fixed, finite number of paths therefore carries essentially all weight: we will now show that these paths are, with overwhelming probability, organised in a "tree-like fashion". Towards this goal, we go back to the original formulation

$$\sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} Z_{r,n}^{e_i} = \sum_{i \le \hat{K}_{\varepsilon}} e^{-n\xi_{p(i)}} \sum_{x \in \mathcal{V}_{r+1,n}^{\rightarrow}: x_2 = e_{p(i)}} \exp\left(-n\sum_{l=2}^{r+1} \xi_{[x]_l}\right) .$$
(3.23)

Note that any directed path of length r + 1 with first step  $(\mathbf{0}, e_i)$ , can only share an edge with another path starting with  $(\mathbf{0}, e_j)$ ,  $i \neq j$  if it goes in the direction  $e_j$  at some point. By this observation for  $i \neq j$  and  $i, j \in \{1, ..., n\}$ 

$$|\{x \in \mathcal{V}_{r+1,n}^{\rightarrow} : x_2 = e_i, \exists x' \in \mathcal{V}_{r+1,n}^{\rightarrow} \text{ s.t. } x'_2 = e_j \text{ and } x \cap x' \neq \emptyset\}| = O(n^{r-1})$$
(3.24)

holds. Combining this fact with the observation

$$\mathbb{E}\exp\left(-n\sum_{l=2}^{r+1}\xi_{[x]_l}\right) = (n+1)^{-r}$$
(3.25)

we see that the total contribution of such paths converges in probability to zero, by Markov inequality, and swapping these intersecting summands for copies of themselves that are independent of paths with different start edge does not change the weak limit. The weak limit of (3.28) therefore coincides with the weak limit of

$$\sum_{i \le \hat{K}_{\varepsilon}} \exp\left(-n\xi_{p(i)}\right) \sum_{x \in \mathcal{V}_{r,n-1}^{\rightarrow}} \exp\left(-n\sum_{l=1}^{r} \xi_{[x]_{l}}^{(p(i))}\right)$$
(3.26)

where  $\xi_{[x]_l}^{(p(i))} = \xi_{[x]_l}$  if  $[x]_l$  cannot be part of a path starting with  $e_{p(j)}$  for some  $j \neq i$  with  $j \leq \hat{K}_{\varepsilon}$ . On the other hand, the  $\xi_{[x]_l}^{(p(i))}$ 's are exponentially distributed and independent of each other for different p(i) and or different  $[x]_l$  as well as independent of all  $(\xi_e)_{e \in \mathbb{E}_n}$ . Finally, we realize that replacing

$$\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_{l}}^{(p(i))}\right) \quad \text{by} \quad \exp\left(-(n-1)\sum_{l=1}^{r}\xi_{[x]_{l}}^{(p(i))}\right) \tag{3.27}$$

causes, by the restriction argument (2.10), an error which vanishes in probability. Collecting all changes and estimates, we have thus shown that the distribution of  $Z_{r+1,n}$  is at most  $2\varepsilon + o_n(1)$ -Prohorov distance away from the weak limit of

$$\sum_{i \le \hat{K}_{\varepsilon}} \exp\left(-n\xi_{p(i)}\right) Z_{r,n-1}^{(i)}, \qquad (3.28)$$

where  $Z_{r,n-1}^{(i)}$ ,  $i \in \mathbb{N}$  are independent copies of  $Z_{r,n-1}$ . By assumption  $Z_{r,n-1}$  converges weakly to  $Z_r$  and by Lemma 6 the smallest finitely many  $n\xi$ 's converge weakly to the first that many points of a PPP( $\mathbb{1}_{\mathbb{R}^+} dx$ ). We conclude that the Prohorov distance of  $Z_{r+1,n}$ and

$$\sum_{i \le \hat{K}_{\varepsilon}} \exp\left(-\hat{\eta}_i\right) Z_r^{(i)},\tag{3.29}$$

is at most by an in *n* vanishing sequence larger than  $2\varepsilon$ . Checking using Markov inequality that the contibution of  $i > \hat{K}_{\varepsilon}$  is vanishing in probability gives that

$$d_P\left(\mathcal{L}(Z_{r+1,n}), (\mathcal{L}Z_{r+1})\right) \to 0 \tag{3.30}$$

has to hold as  $n \to \infty$ . This finishes the induction, and the proof of the n-convergence is thus settled.

We move to the proof of the second claim of Proposition 4, the r-convergence. As mentioned, this will be done via a contraction argument on the space  $\mathcal{P}_2$  of probability measures on  $\mathbb{R}$  with finite second moment. To this end, let  $(\eta_i)_{i\in\mathbb{N}}$  be a  $\operatorname{PPP}(\mathbb{1}_{\mathbb{R}^+}dx)$ . Define  $T: \mathcal{P}_2 \to \mathcal{P}_2$ ,

$$T: \mathcal{P}_2 \to \mathcal{P}_2,$$
  
$$\mu \mapsto \mathcal{L}\left(\sum_{i \in \mathbb{N}} e^{-\eta_i} X_i\right),$$
(3.31)

where  $(X_i)_{i\in\mathbb{N}}$  are independent and identically  $\mu$ -distributed, and independent of  $\eta$ . Note that T is well-defined, i.e., we have that  $T\mu$  has a finite second moment for all  $\mu \in \mathcal{P}_2$  by applying the triangle inequality,  $\mathbb{E}[\sum_{i\in\mathbb{N}} e^{-2\eta_i}] = 1/2$  and independence. Moreover, since  $\mathbb{E}[\sum_{i\in\mathbb{N}} e^{-\eta_i}] = 1$  the map T does not change the first moment. Hence, for the subset

$$\mathcal{P}_{2,1} := \left\{ \mu \in \mathcal{P}_2 : \int x \ d\mu = 1 \right\}$$

the restriction of T to  $\mathcal{P}_{2,1}$  maps to  $\mathcal{P}_{2,1}$ . By construction, it holds that

$$\mathcal{L}(Z_{r+1}) = T\mathcal{L}(Z_r). \tag{3.32}$$

We now endow  $\mathcal{P}_2$  with the minimal  $L_2$ -distance  $\ell_2$ , also called Wasserstein distance of order 2: for  $\mu, \nu \in \mathcal{P}_2$  this is defined by

$$\ell_2(\mu, \nu) = \inf\{\|V - W\|_2 : \mathcal{L}(V) = \mu, \mathcal{L}(W) = \nu\},\$$

where the infimum is over all random variables V, W on a joint probability space with the respective distributions. Convergence in  $\ell_2$  implies weak convergence,  $(\mathcal{P}_2, \ell_2)$  and  $(\mathcal{P}_{2,1}, \ell_2)$  are complete metric spaces. For these topological properties and the existence of optimal couplings used below see, e.g., Ambrosio, Gigli and Savaré [1] or Villani [12]. Within the present setting, in order to prove the r-convergence it suffices to prove that

- The restriction of T to  $\mathcal{P}_{2,1}$  is a strict  $\ell_2$ -contraction.
- The standard exponential distribution is a fixed point of T restricted to  $\mathcal{P}_{2,1}$ .

We remark that T as a map on  $\mathcal{P}_2$  has infinitely many fixed points and that our argument below also implies that these fixed points are exactly the exponential distributions with arbitrary parameter, their negatives, and the Dirac measure in 0. Uniqueness of the fixed point on  $\mathcal{P}_{2,1}$  is immediate by Banach fixed point theorem and the strict contraction property.

Contractivity goes as follows. For  $\mu, \nu \in \mathcal{P}_{2,1}$ , let  $(X_i, Y_i)_{i \in \mathbb{N}}$  be a sequence of independent optimal  $\ell_2$ -couplings, which are also independent of  $\eta$ ; optimal  $\ell_2$ -couplings means here that the pair  $(X_i, Y_i)$  has marginal distributions  $\mu$  and  $\nu$ , and that it attains the infimum in the definition of  $\ell_2$ . It then holds:

$$\ell_2(T\mu, T\nu)^2 \le \mathbb{E}\left[\left(\sum_{i \in \mathbb{N}} e^{-\eta_i} (X_i - Y_i)\right)^2\right].$$
(3.33)

Remark that the off-diagonal terms on the r.h.s. above vanish, since  $X_i - Y_i$  has zero expectation: using this, we thus obtain

$$\ell_2(T\mu, T\nu)^2 \le \mathbb{E}\left[\sum_i e^{-2\eta_i}\right] \mathbb{E}\left[(X_1 - Y_1)^2\right] = \frac{1}{2}\ell_2(\mu, \nu)^2, \qquad (3.34)$$

the last step by optimality of the coupling. This implies that the restriction of the map T to  $\mathcal{P}_{2,1}$  is an  $\ell_2$ -contraction.

It thus remains to prove that the standard exponential distribution is the fixed point of T in  $\mathcal{P}_{2,1}$ . This can be checked via Laplace transformation: consider independent standard exponentials  $X_1, X_2, \ldots$  which are also independent of  $\eta$ . For t > 0,

$$\mathbb{E}\left[\exp\left(-t\sum_{i=1}^{\infty}e^{-\eta_i}X_i\right)\right] = \mathbb{E}\left[\exp\left(-\sum_{i=1}^{\infty}\ln\left(1+te^{-\eta_i}\right)\right)\right]$$
$$= \exp\left(\int_{0}^{\infty}\frac{1}{1+te^{-x}} - 1dx\right) = \frac{1}{1+t},$$
(3.35)

which is the Laplace transform of a standard exponential. This implies ii). The r-convergence therefore immediately follows from Banach fixed point theorem.

## 3.2. Vanishing of the Chen-Stein term. The goal here is to prove (2.17), namely that

$$\lim_{r \to \infty} \lim_{n \to \infty} \operatorname{CS}(r, n) = 0.$$
(3.36)

This requires some additional notation. Let

$$\Sigma_{n,r} \equiv \left\{ (\pi, \pi') \in \Sigma_n \times \Sigma_n : \pi, \pi' \text{ have at least a common edge } e, \\ e = (u, v) \in E, \{ d(u, \mathbf{0}), d(v, \mathbf{0}) \} \in [r, n - r) \right\}.$$

For paths  $(\pi, \pi') \in \Sigma_n \times \Sigma_n$ , we denote by  $\pi \wedge \pi'$  their *overlap*, i.e. the number of edges shared by both paths. Working out the conditional Chen-Stein bound (2.5), we get

$$CS(r,n) = \mathbb{E}d_{TV,\mathcal{F}_{r,n}} \left( \Xi_n(A), \operatorname{Poi} \left( \lambda_n(A) \right) \right)$$

$$\leq \sum_{\pi \in \Sigma_n} \mathbb{E} \left[ \mathbb{E}[I_{\pi}(A) | \mathcal{F}_{r,n}]^2 \right]$$

$$+ \sum_{\star} \mathbb{E} \left[ \mathbb{E}[I_{\pi}(A) | \mathcal{F}_{r,n}] \mathbb{E}[I_{\pi'}(A) | \mathcal{F}_{r,n}] \right]$$

$$+ \sum_{\star} \mathbb{E} \left[ \mathbb{E}[I_{\pi}(A) I_{\pi'}(A) | \mathcal{F}_{r,n}] \right], \qquad (3.37)$$

where  $\sum_{\star}$  denotes summation over all  $(\pi, \pi') \in \Sigma_{n,r} : 1 \leq \pi \wedge \pi' \leq n-2$ . We will prove that all three terms on the r.h.s. of (3.37) vanish in the limit  $n \to \infty$  first, and  $r \to \infty$ next. As the proof is long and technical, we formulate the statements in the form of three Lemmata.

Lemma 8.

$$\lim_{r \to \infty} \lim_{n \to \infty} \sum_{\pi \in \Sigma_n} \mathbb{E} \left[ \mathbb{E} [I_{\pi}(A) | \mathcal{F}_{r,n}]^2 \right] = 0.$$

Lemma 9.

$$\lim_{r \to \infty} \lim_{n \to \infty} \sum_{\star} \mathbb{E} \left[ \mathbb{E} [I_{\pi}(A) | \mathcal{F}_{r,n}] \mathbb{E} [I_{\pi'}(A) | \mathcal{F}_{r,n}] \right] = 0.$$

Lemma 10.

$$\lim_{r \to \infty} \lim_{n \to \infty} \sum_{\star} \mathbb{E} \left[ \mathbb{E} [I_{\pi}(A) I_{\pi'}(A) | \mathcal{F}_{r,n}] \right] = 0.$$

The first contribution is easily taken care of:

Proof of Lemma 8. By symmetry we have that

$$\sum_{\pi \in \Sigma_n} \mathbb{E}[\mathbb{E}[I_{\pi}(A)|\mathcal{F}_{r,n}]^2] = n! \mathbb{E}[\mathbb{E}[I_{\pi^*}(A)|\mathcal{F}_{r,n}]^2], \qquad (3.38)$$

where  $\pi^* \in \Sigma_n$  is arbitrary. It thus follows from the tail-estimates of Lemma 2 that

$$(3.38) = n! \int_{0}^{1+\frac{a}{n}} \left(1 + K\left(1 + \frac{a}{n} - x, n - 2r\right)\right)^{2} \frac{e^{-2\left(1 + \frac{a}{n}\right) + x}\left(1 + \frac{a}{n} - x\right)^{2n - 4r} x^{2r - 1}}{(n - 2r)!^{2}(2r - 1)!} dx$$

$$\leq n! \int_{0}^{1+\frac{a}{n}} \left(1 + e^{\left(1 + \frac{a}{n}\right)} \frac{\left(1 + \frac{a}{n}\right)}{n - 2r}\right)^{2} \frac{e^{-\left(1 + \frac{a}{n}\right)}\left(1 + \frac{a}{n}\right)^{2n - 4r}\left(1 + \frac{a}{n}\right)^{2r - 1}}{(n - 2r)!^{2}(2r - 1)!} dx$$

$$= \frac{n! e^{2a}}{(n - 2r)!^{2}(2r - 1)!} (1 + o_{n}(1)). \qquad (3.39)$$

Since the r.h.s. of (3.39) is vanishing in the large *n*-limit, the proof of Lemma 8 is concluded.

Lemma 9 and 10 require more work. In particular, we will make heavy use of the following combinatorial estimates, which have been established by Fill and Pemantle [8] (see Lemma 2.3, 2.4 and 2.5 p. 598):

**Proposition 11** (Path counting). Let  $\pi'$  be any reference path on the n-dim hypercube connecting **0** and **1**. Denote by f(n,k) the number of paths  $\pi$  that share precisely k edges  $(k \ge 1)$  with  $\pi'$ . Finally, shorten  $\mathfrak{n}_{\mathfrak{e}} \equiv n - 5e(n+3)^{2/3}$ .

• For any K(n) = o(n) as  $n \to \infty$ ,

$$f(n,k) \le (1+o(1))(k+1)(n-k)!$$
(3.40)

uniformly in k for  $k \leq K(n)$ .

• Suppose  $k \leq \mathfrak{n}_{\mathfrak{e}}$ . Then, for n large enough,

$$f(n,k) \le n^6(n-k)!$$
. (3.41)

• Suppose  $k \geq \mathfrak{n}_{\mathfrak{e}}$ . Then, for n large enough,

$$f(n,k) \le (2n^{\frac{7}{8}})^{n-k}(n-k+1).$$
 (3.42)

Proof of Lemma 9. Here and below,  $\kappa_a > 0$  will denote a universal constant not necessarily the same at different occurrences, and which depends solely on a. By symmetry,

$$\sum_{\star} \mathbb{E}[I_{\pi}(A)I_{\pi'}(A)] = n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)]$$
(3.43)

where  $\pi^* \in \Sigma_n$  is arbitrary and  $\sum_{\star,\star}$  standing for summation over

$$\pi' \in \Sigma_n : (\pi^*, \pi') \in \Sigma_{n,r}, 1 \le \pi^* \land \pi' \le n - 2.$$

Let  $k \in \{1, n-2\}$  and  $\pi' \in \Sigma_n, \pi^* \wedge \pi' = k$ . Splitting  $X_{\pi^*}$  and  $X_{\pi'}$  into common/noncommon edges, we obtain

$$\mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] = \mathbb{P}\left(X_{\pi^*} \le 1 + \frac{a}{n}, X_{\pi'} \le 1 + \frac{a}{n}\right) \\ = \int_{\mathbb{R}} \mathbb{P}\left(x + X_{n-k} \le 1 + \frac{a}{n}, x + X'_{n-k} \le 1 + \frac{a}{n} \mid X_k = x\right) \mathbb{P}(X_k \in dx) .$$
(3.44)

In the above,  $X_{n-k}$  and  $X'_{n-k}$  correspond to the compound weights of the non-common edges: these are Gamma(n-k, 1)-distributed random variables;  $X_k$  corresponds to the

weight of the common edges: this is a Gamma(k, 1)-distributed random variable. By construction,  $X_{n-k}, X'_{n-k}$  and  $X_k$  are independent. All in all,

$$\mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] = \int_0^{+\infty} \mathbb{P}\left(x + X_{n-k} \le 1 + \frac{a}{n}\right)^2 \frac{e^{-x}x^{k-1}}{(k-1)!} dx$$

$$\le \frac{\kappa_a}{(n-k)!^2} \int_0^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2(n-k)} \frac{x^{k-1}}{(k-1)!} dx \,. \tag{3.45}$$

The last inequality by the tail-estimate of Lemma 2. Integration by parts then yields

$$\int_{0}^{1+\frac{a}{n}} \left(1+\frac{a}{n}-x\right)^{2(n-k)} x^{k-1} dx \le \kappa_a \frac{(k-1)!(2(n-k))!}{(2n-k)!}.$$
(3.46)

and therefore

$$E[I_{\pi^*}(A)I_{\pi'}(A)] \le \kappa_a \frac{(2(n-k))!}{(2n-k)!(n-k)!^2}.$$
(3.47)

Denoting by f(n, k, r) the number of paths  $\pi'$  that share precisely k edges  $(1 \le k \le n-2)$  with  $\pi^*$  and that satisfy  $(\pi', \pi^*) \in \Sigma_{n,r}$ , we thus have that

$$n! \sum_{\star,\star} \mathbb{E}[I_{\pi^{\star}}(A)I_{\pi'}(A)] = n! \sum_{k=1}^{n-2} f(n,k,r) \mathbb{E}[I_{\pi^{\star}}(A)I_{\pi'}(A)]$$

$$\stackrel{(3.47)}{\leq} \kappa_{a} \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times \frac{n!(2(n-k))!}{(n-k)!(2n-k)!} \qquad (3.48)$$

$$\leq \kappa_{a} \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times \frac{(1-\frac{k}{2n})^{n-k}}{2^{k}(1-\frac{k}{2n})^{2n-k}},$$

the last inequality by Stirling approximation. To lighten notation, remark that with  $\gamma \equiv k/n \in [0, 1]$ , the second factor in the last sum above can be written as

$$\frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}} = \left(\frac{(4(1-\gamma))^{(1-\gamma)}}{(2-\gamma)^{(2-\gamma)}}\right)^n \equiv g(\gamma)^n \,. \tag{3.49}$$

With this, (3.48) takes the form

$$n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] \le \kappa_a \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times g\left(\frac{k}{n}\right)^n.$$
(3.50)

The following observation, whose elementary proof is postponed to the end of this section, will be useful.

**Fact 1.** The function  $g: [0,1] \to \mathbb{R}_+$  defined (3.49) is increasing on [2/3,1). Furthermore,

$$\forall \gamma \le 2/3 : g(\gamma) \le \left(\frac{3}{4}\right)^{\gamma}$$
 (3.51)

In view of Proposition 11, recalling that  $\mathfrak{n}_{\mathfrak{e}} = n - 5e(n+3)^{2/3}$  and with

$$C \equiv \frac{7}{\ln(4/3)},$$
 (3.52)

we split the sum on the r.h.s. of (3.50) into three regimes, to wit:

$$\left(\sum_{k=1}^{C\ln(n)} + \sum_{k=C\ln(n)+1}^{\mathfrak{n}_{\mathfrak{e}}} + \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2}\right) \frac{f(n,k,r)}{(n-k)!} \times g\left(\frac{k}{n}\right)^{n}.$$
(3.53)

Concerning the first sum :

$$\sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n \stackrel{(3.51)}{\leq} \sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!} \left(\frac{3}{4}\right)^k$$

$$\leq \sum_{k=1}^{r-1} \frac{f(n,k,r)}{(n-r+1)!} \left(\frac{3}{4}\right)^k + \sum_{k=r}^{C\ln(n)} \frac{f(n,k)}{(n-k)!} \left(\frac{3}{4}\right)^k \qquad (3.54)$$

$$\leq \sum_{k=1}^{r-1} \frac{f(n,k,r)}{(n-r+1)!} \left(\frac{3}{4}\right)^k + \kappa_a \sum_{k=r}^{C\ln(n)} (k+1) \left(\frac{3}{4}\right)^k ,$$

by Proposition 11.

The function f(n, k, r) counts the number of paths  $\pi'$  that share precisely k edges  $(1 \le k \le n-2)$  with  $\pi^*$  and that satisfy  $(\pi', \pi^*) \in \Sigma_{n,r}$ : we claim that

$$f(n,k,r) \le r!(n-r-1)!n.$$
(3.55)

To see this, recall that the vertices of the hypercube stand in correspondence with the standard basis of  $\mathbb{R}^n$ : every edge is parallel to some unit vector  $e_j$ , where  $e_j$  connects  $(0, \ldots, 0)$  to  $(0, \ldots, 0, 1, 0, \ldots, 0)$  with a 1 in position j. We identify a directed path  $\pi$  from **0** to **1** by a permutation of  $12 \ldots n$ , say  $\pi_1 \pi_2 \ldots \pi_n$ .  $\pi_l$  is giving the direction the path  $\pi$  goes in step l, hence after i steps the path  $\pi_1 \pi_2 \ldots \pi_n$  is at vertex  $\sum_{j \le i} e_{\pi_j}$ . (By a slight abuse of notation,  $\pi_1$  will refer here below to a number between, 1 and n). Let now  $\pi^*$  be the reference path, say  $\pi^* = 12...n$ . We set  $u_i = l$  if the l-th traversed edge by  $\pi'$  is the i-th shared edge of  $\pi'$  and  $\pi^*$ , setting by convention  $r_0 = 0$  and  $r_{k+1} = n + 1$ . Shorten then  $\mathbf{u} \equiv \mathbf{u}(\pi') = (u_0, \ldots, u_{k+1})$ , and  $s_i \equiv u_{i+1} - u_i$ ,  $i = 0, \ldots, k$ . For any sequence  $\mathbf{u}_0 = (u_0, \ldots, u_{k+1})$  with  $0 = u_0 < u_1 < \ldots < u_k < u_{k+1} = n + 1$ , let  $C(\mathbf{u}_0)$  denote the number of paths  $\pi'$  with  $\mathbf{u}(\pi') = \mathbf{u}_0$ . Since the values  $\pi'_{u_i+1}, \ldots, \pi'_{u_i+s_i-1}$  must be a permutation of  $\{u_i + 1, \ldots, u_i + s_i - 1\}$ , one easily sees that  $C(\mathbf{u}) \leq G(\mathbf{u})$ , where

$$G(\mathbf{u}) = \prod_{i=0}^{k} (s_i - 1)!.$$
(3.56)

We also observe that two such paths must have a common edge in the *middle region*  $(\pi', \pi^*) \in \Sigma_{n,r}$ . Let *e* be such an edge: as it turns out, this is quite restrictive. Indeed, it implies that there exists  $u_j \in \{r+1, n-r\}$  for  $j \in \{1, ..., k\}$ . In virtue of (3.56) and log-convexity of factorials, one has at most r!(n-r-1)! paths  $\pi'$  sharing the edge *e* with the reference-path  $\pi^*$ , and at most  $\binom{n}{1} = n$  ways to choose this edge: combining all this settles (3.55).

### FPP IN THE HYPERCUBE

It follows that

$$\sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n \le \sum_{k=1}^{r-1} \frac{r!(n-r-1)!n}{(n-r+1)!} \left(\frac{3}{4}\right)^k + \kappa_a \sum_{k=r}^{+\infty} (k+1) \left(\frac{3}{4}\right)^k.$$
(3.57)

The first sum above clearly tends to 0 as  $n \to \infty$ , whereas the second sum vanishes when  $r \to \infty$ : the first regime in (3.53) therefore yields no contribution in the double limit.

As for the second regime, by Proposition 11,

$$\sum_{k=C\ln(n)}^{\mathbf{n}_{\mathbf{c}}} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^{n} \leq \sum_{k=C\ln(n)}^{\mathbf{n}_{\mathbf{c}}} \frac{f(n,k)}{(n-k)!} g\left(\frac{k}{n}\right)^{n}$$

$$\leq n^{6} \sum_{k=C\ln(n)}^{\mathbf{n}_{\mathbf{c}}} g\left(\frac{k}{n}\right)^{n}$$

$$= n^{6} \left(\sum_{k=C\ln(n)}^{2n/3} g\left(\frac{k}{n}\right)^{n} + \sum_{k=2n/3+1}^{\mathbf{n}_{\mathbf{c}}} g\left(\frac{k}{n}\right)^{n}\right).$$
(3.58)

As pointed out in Fact 1, the g-function is increasing on [2/3, 1), whereas on the "complement" (3.51) holds: these observations, together with (3.58) imply that

$$\sum_{k=C\ln(n)}^{\mathbf{n}_{\mathbf{c}}} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n \le n^6 \left(\sum_{k=C\ln(n)}^{2n/3} \left(\frac{3}{4}\right)^k + \sum_{k=2n/3+1}^{\mathbf{n}_{\mathbf{c}}} g\left(\frac{\mathbf{n}_{\mathbf{c}}}{n}\right)^n\right)$$

$$\le 4n^6 \left(\frac{3}{4}\right)^{C\ln(n)} + n^7 g\left(\frac{\mathbf{n}_{\mathbf{c}}}{n}\right)^n$$

$$= 4 \exp\left\{\left(6 + C\ln(3/4)\right)\ln(n)\right\} + n^7 g\left(\frac{\mathbf{n}_{\mathbf{c}}}{n}\right)^n.$$
(3.59)

In virtue of the choice (3.52) we have that  $6 + C \ln(3/4) = -1$ , hence

$$(3.59) = o_n(1) + n^7 g\left(\frac{\mathbf{n}_{\mathfrak{e}}}{n}\right)^n.$$
(3.60)

By definition of the g-function (3.49) and  $\mathbf{n}_{\mathfrak{e}}$ , it holds:

$$g\left(\frac{\mathfrak{n}_{\mathfrak{e}}}{n}\right)^{n} = \frac{\left(1 - \frac{\mathfrak{n}_{\mathfrak{e}}}{n}\right)^{n-\mathfrak{n}_{\mathfrak{e}}}}{2^{\mathfrak{n}_{\mathfrak{e}}}\left(1 - \frac{\mathfrak{n}_{\mathfrak{e}}}{2n}\right)^{2n-\mathfrak{n}_{\mathfrak{e}}}} = \left(\frac{5e(n+3)^{\frac{2}{3}}}{n}\right)^{5e(n+3)^{\frac{2}{3}}} 2^{10e(n+3)^{\frac{2}{3}}} \left(1 + \frac{5e(n+3)^{\frac{2}{3}}}{n}\right)^{-n-5e(n+3)^{2/3}}.$$
(3.61)

Notice that

$$1 + \frac{5e(n+3)^{\frac{2}{3}}}{n} \ge 1 \text{ and } (n+3)^{\frac{2}{3}} \le 2n^{\frac{2}{3}} \text{ for } n \ge 3,$$
(3.62)

thus

$$(3.61) \le \left(\frac{40e}{n^{1/3}}\right)^{10en^{\frac{2}{3}}} = o(n^{-7}), \qquad (3.63)$$

implying that the second regime in (3.53) yields no contribution in the limit  $n \to +\infty$ .

As for the third, and last regime: by definition of the g-function,

$$\sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n \leq \sum_{k=n_e+1}^{n-2} \frac{f(n,k)}{(n-k)!} \frac{(1-\frac{k}{n})^{n-k}}{2^k (1-\frac{k}{2n})^{2n-k}} \\ \leq \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{(2n^{\frac{7}{8}})^{n-k} (n-k+1)}{(n-k)!} \frac{(1-\frac{k}{n})^{n-k}}{2^k (1-\frac{k}{2n})^{2n-k}},$$
(3.64)

the last step in virtue of Proposition 11. By change of variable,  $n - k \mapsto u$ , we get

$$(3.64) = \sum_{u=2}^{5e(n+3)^{\frac{2}{3}}-1} \left(\frac{8un^{\frac{7}{8}}}{n}\right)^u \frac{(u+1)}{(1+\frac{u}{n})^{n+u}u!} \le \sum_{u=2}^{\infty} \left(\frac{8e}{n^{\frac{1}{8}}}\right)^u (u+1),$$
(3.65)

the last inequality by Stirling's approximation. It thus follows that the contribution of the third and last regime in (3.53) also vanishes as  $n \to +\infty$ . The proof of Lemma 9 is concluded.

We finally provide the elementary

*Proof of Fact 1.* The sign of g' is given by the sign of

$$\frac{d}{d\gamma}\left(\ln(4-4\gamma)(1-\gamma) - \ln(2-\gamma)(2-\gamma)\right) = \ln\left(\frac{2-\gamma}{4-4\gamma}\right).$$

It follows that  $g'(\gamma) \leq 0 \ \forall \gamma \leq 2/3$  and  $g'(\gamma) \geq 0 \ \forall \gamma \geq 2/3$ . Furthermore, since

$$1 - \gamma \le \left(1 - \frac{\gamma}{2}\right)^2,$$

we have

$$g(\gamma) = \frac{(4(1-\gamma))^{(1-\gamma)}}{(2-\gamma)^{(2-\gamma)}} \le (2-\gamma)^{-\gamma} \le \left(\frac{3}{4}\right)^{\gamma}, \qquad (3.66)$$

 $\forall \gamma \leq 2/3$ , settling (3.51).

Proof of Lemma 10. Again by symmetry,

$$\sum_{\star} \mathbb{E}[\mathbb{E}[I_{\pi}(A)|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]]$$

$$= n! \sum_{\star,\star} \mathbb{E}[\mathbb{E}[I_{\pi^{\star}(A)}|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]]$$

$$= n! \sum_{\star,\star} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^{\star}} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\right],$$
(3.67)

where  $\pi^* \in \Sigma_n$  and  $\sum_{\star,\star}$  stands for summation over

 $\pi' \in \Sigma_n, (\pi^*, \pi') \in \Sigma_{n,r} : 1 \le \pi^* \land \pi' \le n - 2.$ 

We split this sum into two parts: the first contribution will stem from paths  $\pi'$  which share less than 2r edges with  $\pi^*$ , in which case  $\pi'$  and  $\pi^*$  are almost independent when n

20

tends to  $+\infty$ ; the second contribution will come from the (fewer) paths which are more correlated with  $\pi^*$ . Precisely, we write:

$$(3.67) = n! \sum_{\star,\star,1} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) \mathbb{P}\left(X_{\pi'} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right)\right] + n! \sum_{\star,\star,2} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) \mathbb{P}\left(X_{\pi'} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right)\right]$$
(3.68)

while  $\sum_{\star,\star,1}$  denotes summation over

$$\pi' \in \Sigma_n, (\pi^*, \pi') \in \Sigma_{n,r} : 1 \le \pi^* \land \pi' \le 2r ,$$

whereas  $\sum_{\star,\star,2}$  stands for summation over

$$\pi' \in \Sigma_n, (\pi^*, \pi') \in \Sigma_{n,r} : 2r+1 \le \pi^* \land \pi' \le n-2.$$

We now proceed to estimate these two sums: in the first case we will exploit the fact that the involved paths are almost independent. To see how this goes, let

$$C_{r,n,\pi'} \equiv \left\{ e = (u,v) \in E_n, \min\{d(u,\mathbf{0}), d(v,\mathbf{0})\} \in [0,r) \cup [n-r,n), \\ e \text{ is a common edge of } \pi' \text{ and } \pi^* \right\},$$

$$(3.69)$$

and denote by  $\#C \equiv |C_{r,n,\pi'}|$  the cardinality of this set. We now make the following observations:

- #C = 0 (i.e.  $C_{r,n,\pi'} = \emptyset$ ) implies that  $\pi'$  and  $\pi^*$  are, conditionally upon  $\mathcal{F}_{r,n}$ , independent.
- If #C > 0, by positivity of exponentials,

$$\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) \leq \mathbb{P}\left(X_{\pi'} - \sum_{e \in C_{r,n,\pi'}} \xi_e \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) \\
= \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right),$$
(3.70)

where  $X_{n-\#C}$  is a Gamma(n - #C, 1)-distributed random variable which is, conditionally upon  $\mathcal{F}_{r,n}$ , independent of  $X_{\pi^*}$ .

Altogether,

$$n! \sum_{\star,\star,1} \mathbb{E} \left[ \mathbb{P} \left( X_{\pi^*} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n} \right) \mathbb{P} \left( X_{\pi'} \le 1 + \frac{a}{n} | \mathcal{F}_{r,n} \right) \right]$$

$$\leq n! \mathbb{P} \left( X_{\pi^*} \le 1 + \frac{a}{n} \right) \sum_{\star,\star,1} \mathbb{P} \left( X_{n-\#C} \le 1 + \frac{a}{n} \right) .$$

$$(3.71)$$

Convergence of the intensity functions (2.1), implies that the first term  $n!\mathbb{P}\left(X_{\pi_1} \leq 1 + \frac{a}{n}\right)$ converges; in particular, it remains bounded as  $n \to \infty$ . It therefore suffices to prove that  $\sum_{\star,\star,1} \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right)$  tends to 0 in the double limit. To see this, denote by f(n,k,r) the number of paths  $\pi'$  that share precisely k edges  $(1 \le k \le n-2)$  with  $\pi^*$  and with  $(\pi', \pi^*) \in \Sigma_{n,r}$ . We then have:

$$\sum_{\star,\star,1} \mathbb{P}\left(X_{n-\#C} \le 1 + \frac{a}{n}\right) = \sum_{k=1}^{2r} f(n,k,r) \mathbb{P}\left(X_{n-\#C} \le 1 + \frac{a}{n}\right)$$

$$\le \sum_{k=1}^{2r} f(n,k) \mathbb{P}\left(X_{n-\#C} \le 1 + \frac{a}{n}\right),$$
(3.72)

where f(n,k) is the number of paths  $\pi'$  that share precisely  $k \ge 1$  edges with  $\pi^*$ . By the tail-estimates from Lemma 2,

$$\mathbb{P}\left(X_{n-\#C} \le 1 + \frac{a}{n}\right) \le \frac{\kappa_a}{(n-\#C)!} \le \frac{\kappa_a}{(n-k+1)!} \,. \tag{3.73}$$

The second inequality holds since two paths in  $\Sigma_{n,r}$  must share an edge in the *complement* of  $C_{r,n,\pi'}$ . Using (3.73) and Proposition 11 we obtain

$$(3.72) \le \kappa_a \sum_{k=1}^{2r} \frac{(n-k)!(k+1)}{(n-k+1)!}, \qquad (3.74)$$

which vanishes as  $n \to \infty$ : the first sum in (3.68) therefore yields a vanishing contribution. As for the second sum, by Cauchy-Schwarz,

$$n! \sum_{\star,\star,2} \mathbb{E}[\mathbb{E}[I_{\pi^{*}(A)}|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]] \leq n! \sum_{\star,\star,2} \mathbb{E}\left[\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)^{2}\right].$$
 (3.75)

By the tail-estimates from Lemma 2, for the expectation on the r.h.s. above it holds

$$\mathbb{E}\left[\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right)^{2}\right] \\
= \int_{0}^{1+\frac{a}{n}} \left(1 + K(1 + \frac{a}{n} - x, n - 2r)\right)^{2} \frac{e^{-2(1+\frac{a}{n})+x}(1 + \frac{a}{n} - x)^{2n-4r}x^{2r-1}}{(n-2r)!^{2}(2r-1)!} dx \quad (3.76) \\
\leq \frac{\kappa_{a}}{(n-2r)!^{2}(2r-1)!} \int_{0}^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2n-4r}x^{2r-1} dx.$$

Integration by parts then yields

$$\int_{0}^{1+\frac{a}{n}} \left(1+\frac{a}{n}-x\right)^{2n-4r} x^{2r-1} dx \qquad \leq \kappa_a \frac{(2n-4r)!(2r-1)!}{(2n-2r)!}, \tag{3.77}$$

Using (3.76) and (3.77) we get

$$(3.75) \le \kappa_a \sum_{k=2r+1}^{n-2} \frac{f(n,k,r)}{(n-2r)!} \frac{n!(2n-4r)!}{(n-2r)!(2n-2r)!} \,. \tag{3.78}$$

It clearly holds that

$$\frac{n!(2n-4r)!}{(n-2r)!(2n-2r)!} \le 1,$$
(3.79)

hence

$$(3.78) \leq \sum_{k=2r+1}^{n-2} \frac{f(n,k,r)}{(n-2r)!} = \left(\sum_{k=2r+1}^{2r+7} + \sum_{k=2r+8}^{n_{\mathfrak{e}}} + \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2}\right) \frac{f(n,k,r)}{(n-2r)!} =: (A) + (B) + (C),$$

$$(3.80)$$

say. By Proposition 11, and worst-case estimates, the following upperbounds hold:

$$(A) \leq \sum_{k=2r+1}^{2r+7} \frac{(k+1)(n-k)!}{(n-2r)!} \leq \kappa_a \frac{7(2r+8)(n-2r-1)!}{(n-2r)!}$$
$$(B) \leq \sum_{k=2r+8}^{\mathfrak{n}_{\mathfrak{e}}} \frac{n^6(n-k)!}{(n-2r)!} \leq n^6 \sum_{k=2r+8}^{\mathfrak{n}_{\mathfrak{e}}} \frac{(n-k)!}{(n-2r)!} \leq n^7 \frac{(n-2r-8)!}{(n-2r)!} \qquad (3.81)$$
$$(C) \leq \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{(2n^{7/8})^{n-k}(n-k+1)}{(n-2r)!} \leq \frac{n^2(2n^{7/8})^{5e(n+3)^{2/3}}}{(n-2r)!}.$$

All three terms are clearly vanishing in the limit  $n \to \infty$ . This implies that the second sum in (3.68) yields no contribution, and the proof of Lemma 10 is thus concluded.

## APPENDIX: THE CONDITIONAL CHEIN-STEIN METHOD

All random variables in the course of the proof are defined on the same probability space  $(\Omega, \mathscr{F}, \mathbb{P})$ . Let  $\mathcal{F} \subset \mathscr{F}$  be a sigma algebra, I is a finite (deterministic) set, and  $(X_i)_{i \in I}$  a family of Bernoulli random variables. We set

$$W \equiv \sum_{i \in I} X_i, \qquad \lambda \equiv \sum_{i \in I} \mathbb{E}(X_i | \mathcal{F}).$$

Since the claim is trivial for  $\lambda = 0$  we assume  $\lambda > 0$  from here onwards. Additionally we denote by  $\widehat{W}$  a random variable which is, conditionally upon  $\mathcal{F}$ , Poi( $\lambda$ )-distributed, i.e.

$$\mathbb{P}(\widehat{W} = k | \mathcal{F})(\omega) = \frac{\lambda(\omega)^k}{k!} e^{-\lambda(\omega)}.$$
(3.82)

(To lighten notation, we will omit henceforth the  $\omega$ -dependence). Assume to be given a bounded,  $\mathcal{F}$ -measurable (possibly random) real-valued function f which satisfies

$$\mathbb{E}(f(W)|\mathcal{F}) = 0$$

and define  $g_f : \mathbb{N} \to \mathbb{R}$  by

$$g_f(0) \equiv 0, \quad g_f(n) \equiv \frac{(n-1)!}{\lambda^n} \sum_{k=0}^{n-1} \frac{f(k)\lambda^k}{k!} \quad n > 0.$$
 (3.83)

We claim that  $g_f$  is  $\mathcal{F}$ -measurable, bounded, and satisfies the following identities:

$$f(n) = \lambda g_f(n+1) - ng_f(n), \qquad n \ge 0,$$
 (3.84)

and

$$g_f(n) = -\frac{(n-1)!}{\lambda^n} \sum_{k=n}^{\infty} \frac{f(k)\lambda^k}{k!} \qquad n > 0.$$
(3.85)

Measurability and first identity follow steadily from the definition. The second identity follows from the fact that  $\mathbb{E}(f(\widehat{W})|\mathcal{F}) = 0$ , whereas boundedness follows from the integral representation of the Taylor rest-term of the exponential function:

$$|g_{f}(n)| \leq \frac{(n-1)! \max_{k \in \mathbb{N}} |f(k)|}{\lambda^{n}} \int_{0}^{\lambda} \frac{t^{n-1}}{(n-1)!} e^{t} dt \leq \frac{\max_{k \in \mathbb{N}} |f(k)| e^{\lambda}}{n}.$$
 (3.86)

Let now  $A \subset \mathbb{N}_0$ , and consider the function

$$f_{A,\lambda}(n) \equiv \mathbb{1}_{n \in A} - \mathbb{P}(\widehat{W} \in A | \mathcal{F}), \quad n \in \mathbb{N}.$$
 (3.87)

This is clearly a bounded,  $\mathcal{F}$ -measurable function which satisfies  $\mathbb{E}(f_{A,\lambda}(\widehat{W})|\mathcal{F}) = 0$ . Therefore, by the above and in particular (3.84), there exists a bounded  $\mathcal{F}$ -measurable function, denoted by  $g_{A,\lambda}$ , which satisfies

$$\mathbb{1}_{n \in A} - \mathbb{P}(\widehat{W} \in A | \mathcal{F}) = \lambda g_{A,\lambda}(n+1) - ng_{A,\lambda}(n), \qquad (3.88)$$

almost surely for any  $n \in \mathbb{N}$ . It follows that

$$\mathbb{1}_{W \in A} - \mathbb{P}(\widehat{W} \in A | \mathcal{F}) = \lambda g_{A,\lambda}(W+1) - W g_{A,\lambda}(W).$$
(3.89)

Taking conditional expectations thus yields

$$\mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widetilde{W} \in A|\mathcal{F}) = \lambda \mathbb{E}(g_{A,\lambda}(W+1)|\mathcal{F}) - \mathbb{E}(Wg_{A,\lambda}(W)|\mathcal{F})$$
$$= \sum_{i \in I} \mathbb{E}(X_i|\mathcal{F})\mathbb{E}(g_{A,\lambda}(W+1)|\mathcal{F}) - \mathbb{E}(X_ig_{A,\lambda}(W)|\mathcal{F}).$$
(3.90)

Consider now the random subset

 $N_i \equiv \{j \in I \setminus \{i\} : X_j \text{ and } X_i \text{ are not conditionally independent given } \mathcal{F}\},$ and denote by  $S^{(i)}$  a random variable which is distributed like  $\sum_{j \in N_i} X_j$  conditionally upon  $\mathcal{F}$  and  $\{X_i = 1\}$ , i.e.

$$\mathbb{P}(S^{(i)} = k|\mathcal{F}) = \mathbb{P}\left(\sum_{j \in N_i} X_j = k, X_i = 1 \middle| \mathcal{F} \right) \middle/ \mathbb{P}(X_i = 1 \middle| \mathcal{F}).$$
(3.91)

if  $\mathbb{P}(X_i = 1 | \mathcal{F}) > 0$ , and arbitrarily defined otherwise. We remark that  $X_i$  and  $(X_j)_{j \in (N_i \cup \{i\})^c}$  are conditionally on  $\mathcal{F}$  independent. Therefore

$$\mathbb{E}(X_i g_{A,\lambda}(W)|\mathcal{F}) = \mathbb{P}(X_i = 1|\mathcal{F})\mathbb{E}\left[g_{A,\lambda}\left(1 + S^{(i)} + \sum_{j \in I \setminus (N_i \cup \{i\})} X_j\right) \middle|\mathcal{F}\right], \quad (3.92)$$

24

$$\mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widehat{W} \in A|\mathcal{F})$$
$$= \sum_{i \in I} \mathbb{E}(X_i|\mathcal{F}) \mathbb{E}\left[g_{A,\lambda}(1+W) - g_{A,\lambda}(1+S^{(i)} + \sum_{j \in I \setminus (N_i \cup \{i\})} X_j) \middle| \mathcal{F}\right].$$
(3.93)

Set now

$$M \equiv \sup\{|g_{A,\lambda}(n+1) - g_{A,\lambda}(n)| : n \in \mathbb{N}_0\}.$$
(3.94)

(Notice that M is  $\mathcal{F}$ -measurable). By the triangle inequality, and worstcase-scenario,

$$\mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) \leq M \sum_{i \in I} \mathbb{E}(X_i|\mathcal{F})\mathbb{E}(X_i + S^{(i)} + \sum_{j \in N_i} X_j|\mathcal{F})$$

$$= M \sum_{i \in I} \left( E(X_i|\mathcal{F})^2 + \sum_{j \in N_i} \left( \mathbb{E}(X_jX_i|\mathcal{F}) + \mathbb{E}(X_j|\mathcal{F})E(X_i|\mathcal{F}) \right) \right).$$
(3.95)

It remains to prove that  $M \leq 1$ . To this end we observe that additivity of  $g_{,\lambda}$  is inherited from  $f_{,\lambda}$ , hence

$$g_{A,\lambda} = \sum_{j \in A} g_{\{j\},\lambda} \,. \tag{3.96}$$

Furthermore,

$$\sum_{j=0}^{\infty} g_{\{j\},\lambda}(n+1) - g_{\{j\},\lambda}(n) = 0, \qquad (3.97)$$

since

$$\sum_{j=0}^{\infty} g_{\{j\},\lambda}(n) \stackrel{(3.96)}{=} g_{\mathbb{N}_0,\lambda}(n) = 0 \quad \forall n \in \mathbb{N},$$
(3.98)

because  $f_{\mathbb{N}_0,\lambda}$  is the zero function. Therefore, for any  $A \subset \mathbb{N}_0$ ,

$$|g_{A,\lambda}(n+1) - g_{A,\lambda}(n)| \le \sum_{j=0}^{\infty} (g_{\{j\},\lambda}(n+1) - g_{\{j\},\lambda}(n))^+.$$
 (3.99)

By (3.83), the definition of f and elementary computations we have, for  $0 < n \leq j$ , that

$$g_{\{j\},\lambda}(n) = -\mathbb{P}(\widehat{W} = j|\mathcal{F}) \sum_{l=0}^{n-1} \frac{(n-1)!}{\lambda^{l+1}(n-1-l)!}.$$
(3.100)

This implies in particular that  $g_{\{j\},\lambda}(n)$  is decreasing in n on [0, j], hence all summands  $j \ge n + 1$  in (3.99) vanish. On the other hand, by (3.85), again the definition of f and elementary computations we have for n > j

$$g_{\{j\},\lambda}(n) = \mathbb{P}(\widehat{W} = j|\mathcal{F}) \sum_{l=0}^{\infty} \frac{\lambda^l (n-1)!}{(n+l)!}.$$
(3.101)

Since this is also decreasing in n, it follows that j = n is the only non-zero summand in (3.99). All in all,

$$M = \sup_{n \in \mathbb{N}} |g_{A,\lambda}(n+1) - g_{A,\lambda}(n)| \le \sup_{n \in \mathbb{N}} |g_{\{n\},\lambda}(n+1) - g_{\{n\},\lambda}(n)|.$$
(3.102)

Now, for n > 0, by (3.100) and (3.101),

$$g_{\{n\},\lambda}(n+1) - g_{\{n\},\lambda}(n) \mid = \\ = \frac{\lambda^n e^{-\lambda}}{n!} \left( \sum_{l=0}^{\infty} \frac{\lambda^l (n-1)!}{(n+l)!} + \sum_{l=0}^{n-1} \frac{(n-1)!}{\lambda^{l+1} (n-1-l)!} \right)$$

$$= \frac{e^{-\lambda}}{n} \left( \sum_{l=n}^{\infty} \frac{\lambda^l}{l!} + \sum_{l=0}^{n-1} \frac{\lambda^l}{l!} \right) = \frac{1}{n} \le 1.$$
(3.103)

On the other hand, for n = 0,

$$|g_{\{0\},\lambda}(1) - g_{\{0\},\lambda}(0)| = \frac{1}{\lambda}(1 - e^{-\lambda}) \le 1, \qquad (3.104)$$

by Taylor estimate. Using (3.103) and (3.104) in (3.102) shows that  $M \leq 1$  as claimed, and concludes the proof of the conditional Chen-Stein method.

#### References

- Ambrosio, Luigi, Gigli, Nicola and Savaré, Giuseppe, Gradient flows in metric spaces and in the space of probability measures, Lectures in Mathematics ETH Zürich, 2nd ed., Birkhäuser Verlag, Basel (2008).
- [2] Bovier, Anton. Gaussian processes on trees: From spin glasses to branching Brownian motion. Cambridge Studies in Advanced Mathematics Vol. 163, Cambridge University Press (2016).
- Bovier, Anton and Lisa Hartung. The extremal process of two-speed branching Brownian motion. Elect. J. Probab. 19, No. 18 (2014): 1-28.
- [4] Bovier, Anton and Lisa Hartung. Variable speed branching Brownian motion 1. Extremal processes in the weak correlation regime. ALEA, Lat. Am. J. Probab. Math. Stat. 12 (2015): 261-291.
- [5] Derrida, Bernard and Herbert Spohn. Polymers on disordered trees, spin glasses, and traveling waves. J. Statist. Phys. 51, no. 5-6 (1988): 817840.
- [6] Fang, Ming and Ofer Zeitouni. Slowdown for time inhomogeneous branching Brownian motion. J. Stat. Phys. 149 no. 1 (2012): 19.
- [7] Fang, Ming and Ofer Zeitouni. Branching random walks in time inhomogeneous environments. Electron. J. Probab. 17 no. 67 (2012): 1-18.
- [8] Fill, James Allen, and Robin Pemantle. Percolation, first-passage percolation and covering times for Richardson's model on the n-cube. The Annals of Applied Probability (1993): 593-629.
- [9] Kistler, Nicola. Derrida's random energy models. From spin glasses to the extremes of correlated radom fields. In: V. Gayrard and N. Kistler (Eds.) Correlated Random Systems: five different methods, Springer Lecture Notes in Mathematics, Vol. 2143 (2015).
- [10] Kallenberg, Olav. Random Measures, Theory and Applications. Springer (2017).
- [11] Kistler, Nicola, Adrien Schertzer and Marius A. Schmidt. First passage percolation in the mean field limit. ArXiv e-prints (2018).
- [12] Villani, Cédric, Optimal transport, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin (2009).

NICOLA KISTLER, J.W. GOETHE-UNIVERSITÄT FRANKFURT, GERMANY. *E-mail address*: kistler@math.uni-frankfurt.de

ADRIEN SCHERTZER, J.W. GOETHE-UNIVERSITÄT FRANKFURT, GERMANY. *E-mail address*: schertzer@math.uni-frankfurt.de

MARIUS A. SCHMIDT, J.W. GOETHE-UNIVERSITÄT FRANKFURT, GERMANY. *E-mail address*: mschmidt@math.uni-frankfurt.de