Submitted to the Annals of Applied Probability
arXiv: arXiv:1808.04598

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

By Nicola Kistler[*], Adrien Schertzer[*] and Marius[†] A. Schmidt

Goethe University Frankfurt a. M. [*] *and University of Basel*[†]

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 $E_n \equiv \{(v, v+e_j) : v, v+e_j \in V_n, j \leq n\}$. 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

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

A graphical rendition is given in Figure 1 below.

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*

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

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*,

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

MSC 2010 subject classifications: 60J80, 60G70, 82B44

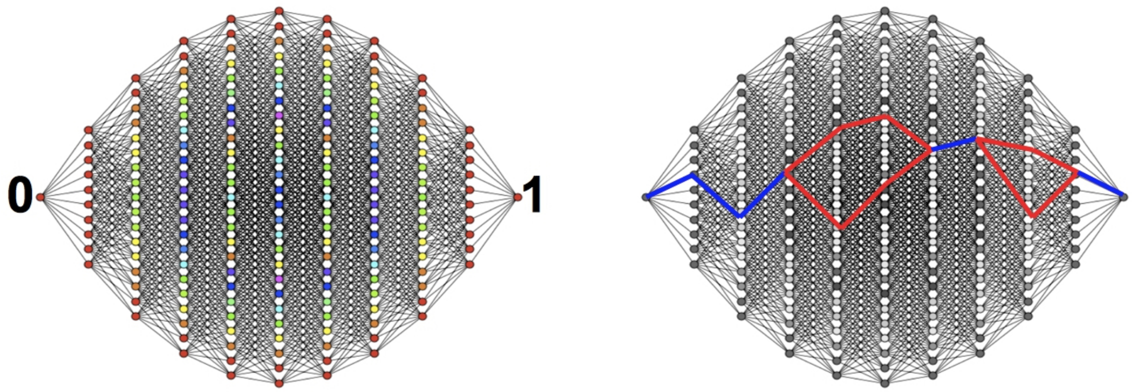

FIG 1. *The 10-dimensional hypercube (left), and two oriented connecting paths (right). Blue edges are common to both paths, whereas paths do not overlap on red edges.*

namely the smallest weight of connecting paths. The limiting value of m_n *to leading order* has been settled by Fill and Pemantle [12], who proved that

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

in probability.

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 [15] an alternative, "modern" approach to (1.3) much inspired by the recent advances in the study of Derrida's random energy models (see [13] 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 Ξ 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

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

It will become clear below, see in particular Remark 6, that the assumption on the distribution of the edge-weights is no restriction, any distribution in the same 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* $\mathbf{0}$ *to* $\mathbf{1}$ *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 [4] 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 [5, 6], see also [9, 10, 11].

The backbone of the proof of Theorem 1 will be presented in Section 2 below. We anticipate that we will check the assumptions of a well-known theorem by Kallenberg by means of the Chein-Stein method [2]. This is arguably the classical route for this type of problems, see e.g. [3, 7, 8, 16]. Contrary to these works, we will however need here a *conditional* version of the Chen-Stein method which we haven't found in the literature, and which may be of independent interest. Section 3 and the Appendix are devoted to the proofs.

2. Strategy of proof. We lighten notation by setting, for $A \subset \mathbb{R}$ a generic subset and π an oriented path,

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

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

- Convergence of the intensity:

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

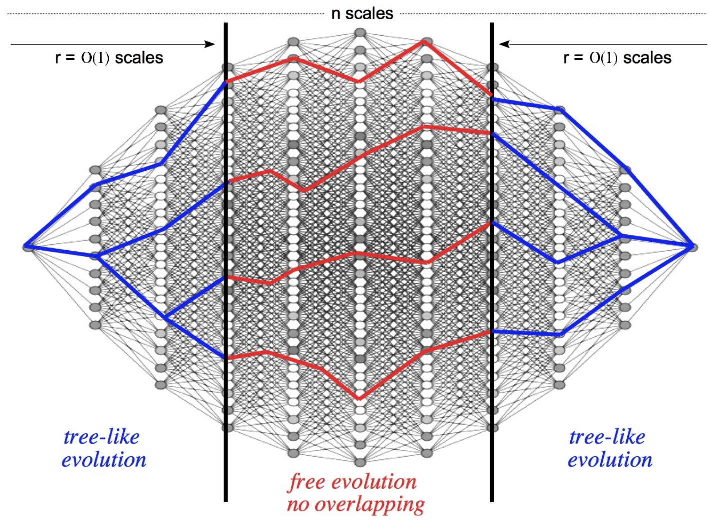

Fig 2. *Four extremal paths. Remark in particular the tree-like evolution close to* **0** *and* **1** *(blue edges) and the (comparatively) longer stretch where paths share no common edge (red). This should be contrasted with the low-dimensional scenario: "loops" in the core of the hypercube, as depicted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

- Convergence of the avoidance function:

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

Theorem 1 then immediately follows in virtue of Kallenberg's Theorem [14, Theorem 4.15].

In the remaining part of this Section we provide a bird's eye view of the main steps involved in the analysis of intensity and avoidance functions. The former 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 [15, 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*

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

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

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 that

$$
\begin{aligned}
\mathbb{E}\Xi_n(A) &= \sum_{\pi\in\Sigma_n} \mathbb{P}\left(n(X_\pi - 1) \le a\right) \\
&= n!\mathbb{P}\left(n(X_{\pi^*} - 1) \le a\right) \qquad \text{(symmetry, } \pi^* \in \Sigma_n \text{ is arbitrary)} \\
&= n!\left\{1 + K\left(1+\frac{a}{n}, n\right)\right\} e^{-1-\frac{a}{n}} \left(\left(1+\frac{a}{n}\right)^+\right)^n (n!)^{-1} \qquad \text{(Lemma 2)} \\
&= (1+o_n(1))e^{-1+a} \\
&= (1+o_n(1))\int_A e^{x-1}dx,
\end{aligned} \tag{2.4}
$$

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, a variant of the classical Chen-Stein method [2] which is tailor-suited to our purposes. Since we haven't found in the literature any similar statement, we provide the proof in the appendix for completeness.

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*

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

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

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

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 $\mathcal{F}$*.*

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, we make the following choices:

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

Proposition 4. *(The double weak-limit). For $\pi_1,..,\pi_{i-1} \in \mathbb{N}$ and $i \le r$, denote by*

$$\left(\eta_{\pi_1,..,\pi_{i-1},\pi_i}\right)_{\pi_i\in\mathbb{N}}, \quad \text{and} \quad \left(\tilde\eta_{\pi_1,..,\pi_{i-1},\pi_i}\right)_{\pi_i\in\mathbb{N}}$$

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

$$Z_r \equiv \sum_{\pi\in\mathbb{N}^r} \exp\left(-\sum_{j=1}^r \eta_{\pi_1\pi_2\ldots\pi_j}\right), \quad \widetilde{Z}_r \equiv \sum_{\pi\in\mathbb{N}^r} \exp\left(-\sum_{j=1}^r \tilde\eta_{\pi_1\pi_2\ldots\pi_j}\right). \tag{2.6}$$

For $A \subset \mathbb{R}$, the following "n-convergence" then holds:

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

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

The proof of the double weak-limit goes via a contraction argument which is given in Section 3.1. Here we shall only point out that both limits Z_r and $\widetilde{Z}_r$ are constructed outgoing from *hierarchical*[1] superpositions of Poisson point processes (PPP for short), and this accounts for

[1]Superpositions of PPP such as those involved in (2.6) are ubiquitous in the Parisi theory of mean field sping glasses, see [13] 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.

the somewhat surprising fact that close to **0** and **1** only tree-like structures contribute to the extremal process in the mean field limit, as depicted in Figure 2.

Most of the technical work will go into the proof of (2.2), which addresses the convergence of avoidance functions. The line of reasoning here goes as follows: recalling that $\Xi_n(A) = \sum_{\pi\in\Sigma_n} I_\pi(A)$, we write

$$
\begin{aligned}
&|\mathbb{P}(\Xi_n(A)=0) - \mathbb{P}(\Xi(A)=0)| = |\mathbb{E}\mathbb{P}(\Xi_n(A)=0 \mid \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A)=0 \mid Z)| \\
&\qquad \le |\mathbb{E}\mathbb{P}(\Xi_n(A)=0 \mid \mathcal{F}_{r,n}) - \mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n})| \\
&\qquad\quad + |\mathbb{E}\mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A)=0 \mid Z)|,
\end{aligned} \tag{2.7}
$$

by the triangle inequality. Furthermore, by convexity,

$$
\begin{aligned}
&|\mathbb{E}\mathbb{P}(\Xi_n(A)=0 \mid \mathcal{F}_{r,n}) - \mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n})| \\
&\quad \le \mathbb{E}\,|\mathbb{P}(\Xi_n(A)=0 \mid \mathcal{F}_{r,n}) - \mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n})| \\
&\quad \le \mathbb{E}d_{TV,\mathcal{F}_{r,n}}(\Xi_n(A), \mathrm{Poi}(\lambda_{r,n}(A))),
\end{aligned} \tag{2.8}
$$

and therefore

$$
\begin{aligned}
|\mathbb{P}(\Xi_n(A)=0) - \mathbb{P}(\Xi(A)=0)| \le &\mathbb{E}d_{TV,\mathcal{F}_{r,n}}(\Xi_n(A), \mathrm{Poi}(\lambda_{r,n}(A))) \\
&+ |\mathbb{E}\mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A)=0 \mid Z)|.
\end{aligned} \tag{2.9}
$$

Concerning the second term on the right-hand side above, it follows from Proposition 4 that

$$
|\mathbb{E}\mathbb{P}(\mathrm{Poi}(\lambda_{r,n}(A))=0 \mid \mathcal{F}_{r,n}) - \mathbb{E}\mathbb{P}(\Xi(A)=0 \mid Z)| = \left|\mathbb{E}\left(e^{-\lambda_{r,n}(A)} - e^{-Z\int_A e^{x-1}dx}\right)\right| \longrightarrow 0, \tag{2.10}
$$

in the double-limit $n\to\infty$ first, and $r\to\infty$ next.

We finally claim that the first term on the right-hand side of (2.9), the "Chen-Stein term", also vanishes in the considered double-limit,

$$
\lim_{r\to\infty}\lim_{n\to\infty}\mathbb{E}d_{TV,\mathcal{F}_{r,n}}(\Xi_n(A), \mathrm{Poi}(\lambda_{r,n}(A))) = 0\,. \tag{2.11}
$$

This is, in fact, the key claim, and its proof is given in Section 3.2 as an application of the conditional Chen-Stein method, Theorem 3. Assuming this for the time being, by combining (2.10) and (2.11), we obtain convergence of the avoidance function and the main Theorem 1 therefore follows.

3. Proofs.

3.1. *The double weak-limit.* The goal of this section is to prove Proposition 4. To see how the limiting objects come about, we lighten notation by setting $V \equiv V_n$, and 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 $\mathbf{0}$ to $\mathbf{1}$ by

$$\begin{aligned}\mathcal{V}_{r,n} =&\{(x,y) \in V^{r+1} \times V^{r+1} : x_1 = \mathbf{0}, d(x_{r+1}, \mathbf{0}) = r, d(y_1, \mathbf{1}) = r, y_{r+1} = \mathbf{1}, \\ &\qquad y_1 - x_{r+1} \in V, (x_i, x_{i+1}), (y_i, y_{i+1}) \in E, \forall i \leq r\}.\end{aligned} \tag{3.1}$$

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

$$\begin{aligned}\Sigma_{x,y} \equiv&\{\pi' \in V^{n-2r+1} : \exists \pi \in \Sigma_n \text{ s.t. } ([\pi]_i)_{i\leq r} = ([x]_i)_{i \leq r} \quad \text{and} \\ &([\pi]_i)_{r<i\leq n-r} = ([\pi']_i)_{r<i\leq n-r}, ([\pi]_i)_{i>n-r} = ([y]_i)_{i>n-r}\}.\end{aligned} \tag{3.2}$$

By definition,

$$\begin{aligned}\lambda_{r,n}(A) &= \sum_{\pi \in \Sigma_n} \mathbb{P}\left(n(X_\pi - 1) \leq a \middle| \mathcal{F}_{r,n}\right) \\ &= \sum_{(x,y)\in\mathcal{V}_{r,n}} \sum_{\pi' \in \Sigma_{x,y}} \mathbb{P}\left(\sum_{i=1}^{n-2r} \xi_{[\pi']_i} \leq 1 + \frac{a}{n} - \sum_{i=1}^{r} \xi_{[x]_i} + \xi_{[y]_i} \middle| \mathcal{F}_{r,n}\right).\end{aligned} \tag{3.3}$$

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 right-hand side of (3.3) equals

$$\sum_{(x,y)\in\mathcal{V}_{r,n}} \left(1 + K(1 + \frac{a}{n} - X_{x,y}, n-2r)\right) \exp\left(-1 - \frac{a}{n} + X_{x,y}\right) \left(\left(1 + \frac{a}{n} - X_{x,y}\right)^+\right)^{n-2r}. \tag{3.4}$$

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

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

for all non-zero summands, and $n \geq a$. 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

$$(3.4) = (1+o_n(1))\left(o_n(1)+e^{-1}\sum_{(x,y)\in\mathcal{V}_{r,n}} \mathbb{1}_{\{X_{x,y}\le\frac{(\ln n)^2}{n}\}}\exp\left((n-2r)\ln\left(1+\frac{a}{n}-X_{x,y}\right)^+\right)\right)$$
$$= (1+o_n(1))\left(o_n(1)+e^{-1+a}\sum_{(x,y)\in\mathcal{V}_{r,n}}\exp\left(-nX_{x,y}\right)\right)$$
$$= (1+o_n(1))\left(o_n(1)+e^{-1+a}\sum_{(x,y)\in\mathcal{V}_{r,n}}\exp -n\sum_{i=1}^{r}\left(\xi_{[x]_i}+\xi_{[y]_i}\right)\right), \tag{3.5}$$

the second step follows by Taylor-expanding the logarithm around 1 to first order, and the third by definition. We now address the sum on the right-hand side of (3.5), on which we perform the aforementioned double limit $n\to\infty$ first and $r\to\infty$ next.

We first address the n-convergence, which states that

$$\lim_{n\to\infty}\sum_{(x,y)\in\mathcal{V}_{r,n}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right) = Z_r\times\widetilde{Z}_r, \tag{3.6}$$

weakly, where $Z_r, \widetilde{Z}_r$ are defined in (2.6). The idea here is to enlarge the set of paths over which the sum is taken, as this enables a useful decoupling, see (3.12) below. Precisely, consider the set of directed paths of length r *from* $\mathbf{0}$,

$$\mathcal{V}^{\leftarrow}_{r,n} = \left\{x\in V^{r+1} : x_1=\mathbf{0}, d(x_{r+1},\mathbf{0})=r, [x]_i\in E, \forall i\le r\right\}, \tag{3.7}$$

and respectively *to* $\mathbf{1}$,

$$\mathcal{V}^{\rightarrow}_{r,n} = \left\{y\in V^{r+1} : y_{r+1}=\mathbf{1}, d(y_1,\mathbf{1})=r, [y]_i\in E, \forall i\le r\right\}. \tag{3.8}$$

One easily checks that

$$\left|\mathcal{V}^{\leftarrow}_{r,n}\times\mathcal{V}^{\rightarrow}_{r,n}\setminus\mathcal{V}_{r,n}\right| = O(n^{2r-1}). \tag{3.9}$$

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

$$\begin{aligned}\sum_{(x,y)\in\mathcal{V}^{\rightarrow}_{r,n}\times\mathcal{V}^{\leftarrow}_{r,n}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right) &=\\ =\sum_{(x,y)\in\mathcal{V}_{r,n}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right)&\\ +\sum_{(x,y)\in(\mathcal{V}^{\rightarrow}_{r,n}\times\mathcal{V}^{\leftarrow}_{r,n})\setminus\mathcal{V}_{r,n}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right)&.\end{aligned} \tag{3.10}$$

and claim that the term on the right-hand side vanishes in probability. Indeed, by a simple computation involving the moment generating function of the exponential distribution, we have

$$\mathbb{E}\left|\sum_{(x,y)\in(\mathcal{V}_{r,n}^{\rightarrow}\times\mathcal{V}_{r,n}^{\leftarrow})\setminus\mathcal{V}_{r,n}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right)\right| = \left|(\mathcal{V}_{r,n}^{\rightarrow}\times\mathcal{V}_{r,n}^{\leftarrow})\setminus\mathcal{V}_{r,n}\right|(n+1)^{-2r} = O\left(\frac{1}{n}\right)\underset{n\to\infty}{\longrightarrow} 0, \tag{3.11}$$

by (3.9). 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,

$$\sum_{(x,y)\in\mathcal{V}_{r,n}^{\rightarrow}\times\mathcal{V}_{r,n}^{\leftarrow}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}+\xi_{[y]_l}\right) = \sum_{x\in\mathcal{V}_{r,n}^{\rightarrow}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}\right)\sum_{y\in\mathcal{V}_{r,n}^{\leftarrow}}\exp\left(-n\sum_{l=1}^{r}\xi_{[y]_l}\right). \tag{3.12}$$

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

$$Z_{r,n}\equiv\sum_{x\in\mathcal{V}_{r,n}^{\rightarrow}}\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}\right)\underset{n\to\infty}{\longrightarrow}\sum_{\pi\in\mathbb{N}^r}\exp\left(\sum_{l=1}^{r}-\eta_{\pi_1\pi_2\ldots\pi_j}\right)\equiv Z_r \tag{3.13}$$

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

LEMMA 5. *Consider $\eta\equiv\sum_{i\in\mathbb{N}}\delta_{\eta_i}$ a $\mathrm{PPP}(\mathbb{1}_{\mathbb{R}^+}dx)$ and independent standard exponentials $(\xi_i)_{i\in\mathbb{N}}$. It then holds:*

$$\sum_{i=1}^{n}\delta_{\xi_i n}\underset{n\to\infty}{\longrightarrow}\eta \tag{3.14}$$

weakly. Furthermore, the following weak limit holds:

$$\sum_{i=1}^{n}\exp(-\xi_i n)\underset{n\to\infty}{\longrightarrow}\sum_{i\in\mathbb{N}}\exp(-\eta_i)\ . \tag{3.15}$$

PROOF OF LEMMA 5. Claim (3.14) is a classical result in extreme value theory. We thus omit its elementary proof. As for the second claim, it is steadily checked (e.g. by Markov's inequality) that the sum on the left-hand side of (3.15) is almost surely finite. In order to prove (3.15) 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

$$\mathbb{E}\exp -t\sum_{i=1}^{n} e^{-\xi_i n} = \mathbb{E}\left(e^{te^{-\xi_1 n}}\right)^n = \left(1+\int_0^{+\infty} e^{-x}(e^{te^{-xn}}-1)dx\right)^n = \left(1+\frac{1}{n}\int_0^{+\infty} e^{-u/n}(e^{te^{-u}}-1)du\right)^n . \tag{3.16}$$

But $e^{-u/n}(e^{te^{-u}}-1) \leq (e^{te^{-u}}-1)$, which is integrable, hence by dominated convergence we have that the right-hand side of (3.16) converges, as $n \uparrow \infty$, to the limit

$$\exp\left(\int_0^{+\infty} (e^{-te^{-x}}-1)dx\right) = \mathbb{E}\exp -t\sum_{i\in\mathbb{N}} e^{-\eta_i}, \tag{3.17}$$

(3.15) is therefore settled. □

Remark 6. *In virtue of Lemma 5, Theorem 1 holds for any choice of edge-weights falling in the same universality class of the exponential distribution, i.e. for which* (3.14) *holds.*

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

$$\mathrm{d_P}\left(\mu,\nu\right) \equiv \inf\left\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon,\ \forall A \subset \mathbb{R}\,\text{closed}\right\} ,$$

where $A^\epsilon \equiv \{x \in \mathbb{R} : \mathrm{d}(A,x) \leq \epsilon\}$ is the ϵ-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 random variables X, Y, slightly abusing notation, one has

$$\mathbb{P}(|X-Y| > \varepsilon) \leq \varepsilon \Rightarrow \mathrm{d_P}(X,Y) \leq \varepsilon\,. \tag{3.18}$$

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

$$Z_{r+1,n} = \sum_{i\leq n} \exp\left(-n\xi_{(\mathbf{0},e_i)}\right) \sum_{x\in\mathcal{V}^{\rightarrow}_{r+1,n}:x_2=e_i} \exp\left(-n\sum_{l=2}^{r+1}\xi_{[x]_l}\right) = \sum_{i\leq n} \exp\left(-n\xi_{(\mathbf{0},e_i)}\right) \times Z^{e_i}_{r,n}\,, \tag{3.19}$$

changing notation for the second sum to lighten exposition.

We claim that it suffices to consider small ξ-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 ε-error in Prohorov distance:

$$\sup_{n,r} \mathrm{d_P}\left(Z_{r+1,n}, \sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}\leq K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i} \right) \leq \varepsilon. \tag{3.20}$$

In fact, for the contribution of large $\xi's$ Markov inequality yields

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}> K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i} > \varepsilon \right) \\
&\qquad\qquad \leq \frac{1}{\varepsilon}\mathbb{E}\left[\sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}> K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i} \right] \\
&\qquad\qquad = \frac{n}{\varepsilon}\mathbb{E}\left[\mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}> K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \right] \times \mathbb{E}\left[Z_{r,n}^{e_i} \right] ,
\end{aligned} \tag{3.21}$$

the last step by independence. One furthermore checks that

$$\mathbb{E}\left[Z_{r,n}^{e_i} \right] = \frac{(n-1)!}{(n-r-1)!}\left(\int_0^\infty e^{-(n+1)x} dx \right)^r = \frac{(n-1)!}{(n-r-1)!}(n+1)^{-r} . \tag{3.22}$$

Thus the right-hand side of (3.21) is *at most*

$$\frac{n}{\varepsilon}\int_{K_\varepsilon/n}^\infty e^{-(n+1)x} dx \times \frac{(n-1)!}{(n-r-1)!}(n+1)^{-r} \leq \frac{\exp -K_\varepsilon}{\varepsilon} = \varepsilon, \tag{3.23}$$

since $K_\varepsilon = -2\ln\varepsilon$. This settles (3.20).

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

$$Z_{r+1,n} \geq \sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}}. \tag{3.24}$$

On the other hand, using that $\mathbb{P}(A) \leq \mathbb{P}(A\cap B) + \mathbb{P}(B^c)$ with obvious identification of the events, we have

$$\begin{aligned}
&\mathbb{P}\left(Z_{r+1,n} \geq \sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}} + \varepsilon\right) \\
&\quad\leq \mathbb{P}\left(Z_{r+1,n} \geq \sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}\leq K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i} + \varepsilon\right) \\
&\qquad + \mathbb{P}\left(\sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} Z_{r,n}^{e_{p(i)}} \leq \sum_{i\leq n} \mathbb{1}_{\{\xi_{(\mathbf{0},e_i)}\leq K_\varepsilon/n\}} e^{-n\xi_{(\mathbf{0},e_i)}} \times Z_{r,n}^{e_i}\right).
\end{aligned} \tag{3.25}$$

While the first term on the right-hand side of (3.25) is at most ε by (3.21) and (3.23), the second term equals

$$\begin{aligned}
&\mathbb{P}\left(\#\{i\leq n : \xi_{(\mathbf{0},e_i)} \leq K_\varepsilon/n\} > \hat{K}_\varepsilon\right) \\
&\qquad \leq n\mathbb{P}(\xi_{(\mathbf{0},e_1)} \leq K_\varepsilon/n)/\hat{K}_\varepsilon \leq K_\varepsilon/\hat{K}_\varepsilon \leq \varepsilon\,,
\end{aligned} \tag{3.26}$$

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

All in all, in virtue of (3.18), the above considerations imply that

$$\sup_{n,r} \mathrm{d_P}\left(Z_{r+1,n}, \sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} Z_{r,n}^{e_p(i)}\right) \leq 2\varepsilon\,. \tag{3.27}$$

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

$$\sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} Z_{r,n}^{e_i} = \sum_{i\leq \hat{K}_\varepsilon} e^{-n\xi_{p(i)}} \sum_{x\in \mathcal{V}^{\rightarrow}_{r+1,n}: x_2=e_{p(i)}} \exp\left(-n\sum_{l=2}^{r+1} \xi_{[x]_l}\right)\,. \tag{3.28}$$

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 $i, j \in \{1, .., n\}$

$$|\{x \in \mathcal{V}^{\rightarrow}_{r+1,n} : x_2 = e_i, \exists x' \in \mathcal{V}^{\rightarrow}_{r+1,n} \text{ s.t. } x'_2 = e_j \text{ and } x\cap x' \neq \emptyset\}| = O(n^{r-1}) \tag{3.29}$$

holds. Combining this fact with the observation

$$\mathbb{E}\exp\left(-n\sum_{l=2}^{r+1} \xi_{[x]_l}\right) = (n+1)^{-r} \tag{3.30}$$

we see that the total contribution of such paths converges in probability to zero, by Markov inequality, and swapping the intersecting summands for copies that are independent of paths with different start edge does not change the weak limit. The weak limit of (3.33) therefore coincides with the weak limit of

$$\sum_{i\leq \hat{K}_\varepsilon} \exp\left(-n\xi_{p(i)}\right) \sum_{x\in\mathcal{V}_{r,n-1}^{\rightarrow}} \exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}^{(p(i))}\right) \tag{3.31}$$

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

$$\exp\left(-n\sum_{l=1}^{r}\xi_{[x]_l}^{(p(i))}\right) \quad \text{by} \quad \exp\left(-(n-1)\sum_{l=1}^{r}\xi_{[x]_l}^{(p(i))}\right) \tag{3.32}$$

causes, by the restriction argument (3.5), 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

$$\sum_{i\leq \hat{K}_\varepsilon} \exp\left(-n\xi_{p(i)}\right) Z_{r,n-1}^{(i)}, \tag{3.33}$$

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 5 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

$$\sum_{i\leq \hat{K}_\varepsilon} \exp\left(-\hat{\eta}_i\right) Z_r^{(i)}, \tag{3.34}$$

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

$$d_P\left(\mathcal{L}(Z_{r+1,n}), \mathcal{L}(Z_{r+1})\right) \to 0 \tag{3.35}$$

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. 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, denote again by $(\eta_i)_{i\in\mathbb{N}}$ a PPP$(\mathbb{1}_{\mathbb{R}^+}dx)$. Define

$$\begin{aligned} T : \mathcal{P}_2 &\to \mathcal{P}_2, \\ \mu &\mapsto \mathcal{L}\left(\sum_{i\in\mathbb{N}} e^{-\eta_i} X_i\right), \end{aligned} \tag{3.36}$$

where $(X_i)_{i\in\mathbb{N}}$ are independent and identically μ-distributed, and independent of η. 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

$$\mathcal{L}(Z_{r+1}) = T\mathcal{L}(Z_r). \tag{3.37}$$

We now endow $\mathcal{P}_2$ with the minimal L_2-distance ℓ_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 taken over all probability distributions on $\mathcal{P} \times \mathcal{P}$ whose marginals are μ and ν respectively. Convergence in ℓ_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 [18]. 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 ℓ_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* ℓ_2-couplings, which are also independent of η; optimal ℓ_2-couplings means here that the pair (X_i, Y_i) has marginal distributions μ and ν, and that it attains the infimum in the definition of ℓ_2. It then holds:

$$\ell_2(T\mu, T\nu)^2 \le \mathbb{E}\left[\left(\sum_{i\in\mathbb{N}} e^{-\eta_i}(X_i - Y_i)\right)^2\right]. \tag{3.38}$$

Remark that the off-diagonal terms on the right-hand side above vanish, since $X_i - Y_i$ has zero expectation. Using this, we thus obtain

$$\ell_2(T\mu, T\nu)^2 \leq \mathbb{E}\left[\sum_i e^{-2\eta_i}\right] \mathbb{E}\left[(X_1 - Y_1)^2\right] = \frac{1}{2}\ell_2(\mu,\nu)^2, \tag{3.39}$$

the last step follows by optimality of the coupling. This implies that the restriction of the map T to $\mathcal{P}_{2,1}$ is an ℓ_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 $X_1, X_2, ...$ which are also independent of η. For $t > 0$,

$$\begin{aligned}\mathbb{E}\left[\exp\left(-t\sum_{i=1}^{\infty} e^{-\eta_i} X_i\right)\right] &= \mathbb{E}\left[\exp\left(-\sum_{i=1}^{\infty} \ln\left(1 + te^{-\eta_i}\right)\right)\right] \\ &= \exp\left(\int_0^{\infty} \frac{1}{1+te^{-x}} - 1dx\right) = \frac{1}{1+t},\end{aligned} \tag{3.40}$$

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 (2.11), namely that

$$\lim_{r\to\infty}\lim_{n\to\infty} \mathbb{E}d_{TV,\mathcal{F}_{r,n}}\left(\Xi_n(A), \mathrm{Poi}\left(\lambda_{r,n}(A)\right)\right) = 0\,. \tag{3.41}$$

This requires some additional notation. Let

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

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

$$\begin{aligned}\mathbb{E}d_{TV,\mathcal{F}_{r,n}}\left(\Xi_n(A), \mathrm{Poi}\left(\lambda_n(A)\right)\right) \leq &\sum_{\pi\in\Sigma_n} \mathbb{E}\left[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]^2\right] \\ &+ \sum_{\star} \mathbb{E}\left[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]\right] \\ &+ \sum_{\star} \mathbb{E}\left[\mathbb{E}[I_\pi(A)I_{\pi'}(A)|\mathcal{F}_{r,n}]\right],\end{aligned} \tag{3.42}$$

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 right-hand side of (3.42) 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 7.
$$\lim_{r\to\infty}\lim_{n\to\infty}\sum_{\pi\in\Sigma_n}\mathbb{E}\left[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]^2\right] = 0\,.$$

Lemma 8.
$$\lim_{r\to\infty}\lim_{n\to\infty}\sum_{\star}\mathbb{E}\left[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]\right] = 0\,.$$

Lemma 9.
$$\lim_{r\to\infty}\lim_{n\to\infty}\sum_{\star}\mathbb{E}\left[\mathbb{E}[I_\pi(A)I_{\pi'}(A)|\mathcal{F}_{r,n}]\right] = 0\,.$$

The first contribution is easily taken care of:

Proof of Lemma 7. By symmetry we have that

$$\sum_{\pi\in\Sigma_n}\mathbb{E}[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]^2] = n!\mathbb{E}[\mathbb{E}[I_{\pi^*}(A)|\mathcal{F}_{r,n}]^2]\,, \tag{3.43}$$

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

$$\begin{aligned}(3.43) &= n!\int_0^{1+\frac{a}{n}}\Big(1+K(1+\frac{a}{n}-x, n-2r)\Big)^2\frac{e^{-2(1+\frac{a}{n})+x}(1+\frac{a}{n}-x)^{2n-4r}x^{2r-1}}{(n-2r)!^2(2r-1)!}dx\\ &\leq n!\int_0^{1+\frac{a}{n}}\left(1+e^{(1+\frac{a}{n})}\frac{(1+\frac{a}{n})}{n-2r}\right)^2\frac{e^{-(1+\frac{a}{n})}(1+\frac{a}{n})^{2n-4r}(1+\frac{a}{n})^{2r-1}}{(n-2r)!^2(2r-1)!}dx\\ &= \frac{n!e^{2a}}{(n-2r)!^2(2r-1)!}(1+o_n(1))\,.\end{aligned} \tag{3.44}$$

Since the right-hand side of (3.44) is vanishing in the large n-limit, the proof of Lemma 7 is concluded. □

Lemma 8 and 9 require more work. In particular, we will make heavy use of the following combinatorial estimates, which have been established by Fill and Pemantle [12] (see Lemma 2.3, 2.4 and 2.5 p. 598):

Proposition 10 (Path counting). *Let π' be any reference path on the n-dim hypercube connecting $\mathbf{0}$ and $\mathbf{1}$. Denote by $f(n,k)$ the number of paths π that share precisely k edges ($k \geq 1$) with π'. Finally, shorten $\mathfrak{n}_\mathfrak{e} \equiv n - 5e(n+3)^{2/3}$.*

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

$$f(n,k) \le (1+o(1))(k+1)(n-k)! \tag{3.45}$$

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

- *Suppose $k \le \mathfrak{n}_\mathfrak{c}$. Then, for n large enough,*

$$f(n,k) \le n^6 (n-k)!\,. \tag{3.46}$$

- *Suppose $k \ge \mathfrak{n}_\mathfrak{c}$. Then, for n large enough,*

$$f(n,k) \le (2n^{\frac{7}{8}})^{n-k}(n-k+1)\,. \tag{3.47}$$

Proof of Lemma 8. Here and below, $\kappa_a > 0$ will denote a universal constant not necessarily the same at different occurences, and which depends solely on a. By symmetry,

$$\sum_{\star} \mathbb{E}[I_\pi(A) I_{\pi'}(A)] = n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A) I_{\pi'}(A)] \tag{3.48}$$

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

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

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

$$\begin{aligned}\mathbb{E}[I_{\pi^*}(A) I_{\pi'}(A)] &= \mathbb{P}\left(X_{\pi^*} \le 1 + \frac{a}{n}, X_{\pi'} \le 1 + \frac{a}{n}\right) \\ &= \int_{\mathbb{R}} \mathbb{P}\left(x + X_{n-k} \le 1 + \frac{a}{n}, x + X'_{n-k} \le 1 + \frac{a}{n} \mid X_k = x\right) \mathbb{P}(X_k \in dx)\,.\end{aligned} \tag{3.49}$$

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, with the tail-estimate of Lemma 2, one has

$$\begin{aligned}\mathbb{E}[I_{\pi^*}(A) I_{\pi'}(A)] &= \int_0^{+\infty} \mathbb{P}\left(x + X_{n-k} \le 1 + \frac{a}{n}\right)^2 \frac{e^{-x} x^{k-1}}{(k-1)!} dx \\ &\le \frac{\kappa_a}{(n-k)!^2} \int_0^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2(n-k)} \frac{x^{k-1}}{(k-1)!} dx\,.\end{aligned} \tag{3.50}$$

Integration by parts then yields

$$\int_0^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2(n-k)} x^{k-1} dx \le \kappa_a \frac{(k-1)!(2(n-k))!}{(2n-k)!}\,. \tag{3.51}$$

and therefore

$$E[I_{\pi^*}(A)I_{\pi'}(A)] \leq \kappa_a \frac{(2(n-k))!}{(2n-k)!(n-k)!^2}. \tag{3.52}$$

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

$$\begin{aligned} n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] &= n! \sum_{k=1}^{n-2} f(n,k,r)\mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] \\ &\overset{(3.54)}{\leq} \kappa_a \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times \frac{n!(2(n-k))!}{(n-k)!(2n-k)!} \end{aligned} \tag{3.53}$$

hence

$$n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] \leq \kappa_a \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times \frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}}, \tag{3.54}$$

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

$$\frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}} = \left(\frac{(4(1-\gamma))^{(1-\gamma)}}{(2-\gamma)^{(2-\gamma)}} \right)^n \equiv g(\gamma)^n \,. \tag{3.55}$$

With this, (3.53) takes the form

$$n! \sum_{\star,\star} \mathbb{E}[I_{\pi^*}(A)I_{\pi'}(A)] \leq \kappa_a \sum_{k=1}^{n-2} \frac{f(n,k,r)}{(n-k)!} \times g\left(\frac{k}{n}\right)^n. \tag{3.56}$$

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 in (3.55) is increasing on $[2/3, 1)$. Furthermore,*

$$\forall \gamma \leq 2/3 : \; g(\gamma) \leq \left(\frac{3}{4}\right)^\gamma \,. \tag{3.57}$$

In view of Proposition 10, recalling that $\mathfrak{n}_\mathfrak{c} = n - 5e(n+3)^{2/3}$ and with

$$C \equiv \frac{7}{\ln{(4/3)}}\,, \tag{3.58}$$

we split the sum on the right-hand side of (3.56) into three regimes, to wit:

$$\left(\sum_{k=1}^{C\ln(n)} + \sum_{k=C\ln(n)+1}^{\mathfrak{n}_\mathfrak{e}} + \sum_{k=\mathfrak{n}_\mathfrak{e}+1}^{n-2}\right) \frac{f(n,k,r)}{(n-k)!} \times g\left(\frac{k}{n}\right)^n . \tag{3.59}$$

Concerning the first sum, by (3.57) it holds

$$\begin{aligned}\sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n &\leq \sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!}\left(\frac{3}{4}\right)^k \\ &\leq \sum_{k=1}^{r-1} \frac{f(n,k,r)}{(n-r+1)!}\left(\frac{3}{4}\right)^k + \sum_{k=r}^{C\ln(n)} \frac{f(n,k)}{(n-k)!}\left(\frac{3}{4}\right)^k \\ &\leq \sum_{k=1}^{r-1} \frac{f(n,k,r)}{(n-r+1)!}\left(\frac{3}{4}\right)^k + \kappa_a \sum_{k=r}^{C\ln(n)} (k+1)\left(\frac{3}{4}\right)^k ,\end{aligned} \tag{3.60}$$

by Proposition 10.

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

$$f(n,k,r) \leq r!(n-r-1)!n. \tag{3.61}$$

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 $(0,\dots,0)$ to $(0,\dots,0,1,0,\dots,0)$ with a 1 in position j. We identify a directed path π from $\mathbf{0}$ to $\mathbf{1}$ by a permutation of $12\dots n$, say $\pi_1\pi_2\dots\pi_n$. π_l is giving the direction the path π goes in step l, hence after i steps the path $\pi_1\pi_2\dots\pi_n$ is at vertex $\sum_{j\leq i} e_{\pi_j}$. (By a slight abuse of notation, π_1 will refer here below to a number between, 1 and n). Let now π^* be the reference path, say $\pi^* = 12...n$. We set $u_i = l$ if the l-th traversed edge by π' is the i-th shared edge of π' and π^*, setting by convention $r_0 = 0$ and $r_{k+1} = n+1$. Shorten then $\mathbf{u} \equiv \mathbf{u}(\pi') = (u_0,...,u_{k+1})$, and $s_i \equiv u_{i+1} - u_i$, $i = 0,...,k$. For any sequence $\mathbf{u}_0 = (u_0,...,u_{k+1})$ with $0 = u_0 < u_1 < ... < u_k < u_{k+1} = n+1$, let $C(\mathbf{u}_0)$ denote the number of paths π' with $\mathbf{u}(\pi') = \mathbf{u}_0$. Since the values $\pi'_{u_i+1},...,\pi'_{u_i+s_i-1}$ must be a permutation of $\{u_i+1,...,u_i+s_i-1\}$, one easily sees that $C(\mathbf{u}) \leq G(\mathbf{u})$, where

$$G(\mathbf{u}) = \prod_{i=0}^{k} (s_i - 1)! . \tag{3.62}$$

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 (3.62) and log-convexity of factorials, one has at most $r!(n-r-1)!$ paths π' sharing the edge e with the reference-path π^*, and at most $\binom{n}{1} = n$ ways to choose this edge: combining all this settles (3.61).

It follows that

$$\sum_{k=1}^{C\ln(n)} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n \leq \sum_{k=1}^{r-1} \frac{r!(n-r-1)!n}{(n-r+1)!}\left(\frac{3}{4}\right)^k + \kappa_a \sum_{k=r}^{+\infty}(k+1)\left(\frac{3}{4}\right)^k. \tag{3.63}$$

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 (3.59) therefore yields no contribution in the double limit.

As for the second regime, by Proposition 10,

$$\begin{aligned}
\sum_{k=C\ln(n)}^{\mathfrak{n}_\mathfrak{e}} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n &\leq \sum_{k=C\ln(n)}^{\mathfrak{n}_\mathfrak{e}} \frac{f(n,k)}{(n-k)!} g\left(\frac{k}{n}\right)^n \\
&\leq n^6 \sum_{k=C\ln(n)}^{\mathfrak{n}_\mathfrak{e}} g\left(\frac{k}{n}\right)^n \\
&= n^6 \left(\sum_{k=C\ln(n)}^{2n/3} g\left(\frac{k}{n}\right)^n + \sum_{k=2n/3+1}^{\mathfrak{n}_\mathfrak{e}} g\left(\frac{k}{n}\right)^n \right).
\end{aligned} \tag{3.64}$$

As pointed out in Fact 1, the g-function is increasing on $[2/3, 1)$, whereas on the "complement" (3.57) holds: these observations, together with (3.64) imply that

$$\begin{aligned}
\sum_{k=C\ln(n)}^{\mathfrak{n}_\mathfrak{e}} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n &\leq n^6 \left(\sum_{k=C\ln(n)}^{2n/3} \left(\frac{3}{4}\right)^k + \sum_{k=2n/3+1}^{\mathfrak{n}_\mathfrak{e}} g\left(\frac{\mathfrak{n}_\mathfrak{e}}{n}\right)^n \right) \\
&\leq 4n^6 \left(\frac{3}{4}\right)^{C\ln(n)} + n^7 g\left(\frac{\mathfrak{n}_\mathfrak{e}}{n}\right)^n \\
&= 4\exp\left\{(6 + C\ln(3/4))\ln(n)\right\} + n^7 g\left(\frac{\mathfrak{n}_\mathfrak{e}}{n}\right)^n.
\end{aligned} \tag{3.65}$$

In virtue of the choice (3.58) we have that $6 + C\ln(3/4) = -1$, hence

$$(3.65) = o_n(1) + n^7 g\left(\frac{\mathfrak{n}_\mathfrak{e}}{n}\right)^n. \tag{3.66}$$

By definition of the g-function (3.55) and $\mathfrak{n}_\mathfrak{e}$, it holds:

$$\begin{aligned}
g\left(\frac{\mathfrak{n}_\mathfrak{e}}{n}\right)^n &= \frac{(1-\frac{\mathfrak{n}_\mathfrak{e}}{n})^{n-\mathfrak{n}_\mathfrak{e}}}{2^{\mathfrak{n}_\mathfrak{e}}(1-\frac{\mathfrak{n}_\mathfrak{e}}{2n})^{2n-\mathfrak{n}_\mathfrak{e}}} \\
&= \left(\frac{5e(n+3)^{\frac{2}{3}}}{n}\right)^{5e(n+3)^{\frac{2}{3}}} 2^{10e(n+3)^{\frac{2}{3}}} \left(1+\frac{5e(n+3)^{\frac{2}{3}}}{n}\right)^{-n-5e(n+3)^{2/3}}.
\end{aligned} \tag{3.67}$$

Notice that

$$1+\frac{5e(n+3)^{\frac{2}{3}}}{n} \geq 1 \text{ and } (n+3)^{\frac{2}{3}} \leq 2n^{\frac{2}{3}} \text{ for } n \geq 3, \tag{3.68}$$

thus

$$g\left(\frac{\mathfrak{n}_{\mathfrak{e}}}{n}\right)^n \leq \left(\frac{40e}{n^{1/3}}\right)^{10en^{\frac{2}{3}}} = o(n^{-7}), \tag{3.69}$$

implying that the second regime in (3.59) yields no contribution in the limit $n \to +\infty$.

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

$$\begin{aligned}\sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{f(n,k,r)}{(n-k)!} g\left(\frac{k}{n}\right)^n &\leq \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{f(n,k)}{(n-k)!} \frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}} \\ &\leq \sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{(2n^{\frac{7}{8}})^{n-k}(n-k+1)}{(n-k)!} \frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}},\end{aligned} \tag{3.70}$$

the last step in virtue of Proposition 10. By change of variable, $n-k \mapsto u$, we get

$$\begin{aligned}\sum_{k=\mathfrak{n}_{\mathfrak{e}}+1}^{n-2} \frac{(2n^{\frac{7}{8}})^{n-k}(n-k+1)}{(n-k)!} \frac{(1-\frac{k}{n})^{n-k}}{2^k(1-\frac{k}{2n})^{2n-k}} &= \sum_{u=2}^{5e(n+3)^{\frac{2}{3}}-1} \left(\frac{8un^{\frac{7}{8}}}{n}\right)^u \frac{(u+1)}{(1+\frac{u}{n})^{n+u}u!} \\ &\leq \sum_{u=2}^{\infty} \left(\frac{8e}{n^{\frac{1}{8}}}\right)^u (u+1)\end{aligned} \tag{3.71}$$

by Stirling's approximation. It thus follows that the contribution of the third and last regime in (3.59) also vanishes as $n \to +\infty$. The proof of Lemma 8 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

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

we have

$$g(\gamma) = \frac{(4(1-\gamma))^{(1-\gamma)}}{(2-\gamma)^{(2-\gamma)}} \leq (2-\gamma)^{-\gamma} \leq \left(\frac{3}{4}\right)^{\gamma}, \tag{3.72}$$

$\forall \gamma \leq 2/3$, settling (3.57). □

Proof of Lemma 9. Again by symmetry,

$$\begin{aligned}
&\sum_{\star} \mathbb{E}[\mathbb{E}[I_\pi(A)|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]] \\
&\qquad = n! \sum_{\star,\star} \mathbb{E}[\mathbb{E}[I_{\pi^*(A)}|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]] \\
&\qquad = n! \sum_{\star,\star} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\right],
\end{aligned} \tag{3.73}$$

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

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

We split this sum into two parts, the first contribution will stem from paths π' which share less than $2r$ edges with π^*, in which case π' and π^* are almost independent when n tends to $+\infty$; the second contribution will come from the (fewer) paths which are more correlated with π^*. Precisely, we write

$$\begin{aligned}
(3.73) =& n! \sum_{\star,\star,1} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\right] \\
&+ n! \sum_{\star,\star,2} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)\right]
\end{aligned} \tag{3.74}$$

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

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

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

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

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

$$\begin{aligned}
C_{r,n,\pi'} \equiv \Big\{ & e = (u,v) \in E_n, \min\{d(u,\mathbf{0}), d(v,\mathbf{0})\} \in [0,r) \cup [n-r,n), \\
& \text{e is a common edge of } \pi' \text{ and } \pi^* \Big\},
\end{aligned} \tag{3.75}$$

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 π' and π^* are, conditionally upon $\mathcal{F}_{r,n}$, independent.
- If $\#C > 0$, by positivity of exponentials,

$$\begin{aligned}\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) &\leq \mathbb{P}\left(X_{\pi'} - \sum_{e \in C_{r,n,\pi'}} \xi_e \leq 1 + \frac{a}{n} \middle| \mathcal{F}_{r,n}\right) \\ &= \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n} \middle| \mathcal{F}_{r,n}\right),\end{aligned} \tag{3.76}$$

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

Altogether,

$$\begin{aligned}n! \sum_{\star,\star,1} \mathbb{E}\left[\mathbb{P}\left(X_{\pi^*} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right) \mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n} | \mathcal{F}_{r,n}\right)\right] \\ \leq n! \mathbb{P}\left(X_{\pi^*} \leq 1 + \frac{a}{n}\right) \sum_{\star,\star,1} \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right).\end{aligned} \tag{3.77}$$

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 π' that share precisely k edges $(1 \leq k \leq n-2)$ with π^* and with $(\pi', \pi^*) \in \Sigma_{n,r}$. We then have

$$\begin{aligned}\sum_{\star,\star,1} \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right) &= \sum_{k=1}^{2r} f(n, k, r) \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right) \\ &\leq \sum_{k=1}^{2r} f(n, k) \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right),\end{aligned} \tag{3.78}$$

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

$$\mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right) \leq \frac{\kappa_a}{(n - \#C)!} \leq \frac{\kappa_a}{(n - k + 1)!}. \tag{3.79}$$

The second inequality holds since two paths in $\Sigma_{n,r}$ must share an edge in the *complement* of $C_{r,n,\pi'}$. Using (3.79) and Proposition 10 we obtain

$$\sum_{\star,\star,1} \mathbb{P}\left(X_{n-\#C} \leq 1 + \frac{a}{n}\right) \leq \kappa_a \sum_{k=1}^{2r} \frac{(n-k)!(k+1)}{(n-k+1)!}, \tag{3.80}$$

which vanishes as $n \to \infty$, the first sum in (3.74) therefore yields a vanishing contribution. As for the second sum, by Cauchy-Schwarz,

$$n! \sum_{\star,\star,2} \mathbb{E}[\mathbb{E}[I_{\pi^*(A)}|\mathcal{F}_{r,n}]\mathbb{E}[I_{\pi'}(A)|\mathcal{F}_{r,n}]] \leq n! \sum_{\star,\star,2} \mathbb{E}\left[\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)^2\right]. \tag{3.81}$$

By the tail-estimates from Lemma 2, for the expectation on the right-hand side above it holds

$$\begin{aligned} &\mathbb{E}\left[\mathbb{P}\left(X_{\pi'} \leq 1 + \frac{a}{n}|\mathcal{F}_{r,n}\right)^2\right] \\ &= \int_0^{1+\frac{a}{n}} \left(1 + K(1 + \frac{a}{n} - x, n-2r)\right)^2 \frac{e^{-2(1+\frac{a}{n})+x}(1+\frac{a}{n}-x)^{2n-4r}x^{2r-1}}{(n-2r)!^2(2r-1)!} dx \\ &\leq \frac{\kappa_a}{(n-2r)!^2(2r-1)!} \int_0^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2n-4r} x^{2r-1} dx. \end{aligned} \tag{3.82}$$

Integration by parts then yields

$$\int_0^{1+\frac{a}{n}} \left(1 + \frac{a}{n} - x\right)^{2n-4r} x^{2r-1} dx \qquad \leq \kappa_a \frac{(2n-4r)!(2r-1)!}{(2n-2r)!}, \tag{3.83}$$

Using (3.82) and (3.83) we get

$$(3.81) \leq \kappa_a \sum_{k=2r+1}^{n-2} \frac{f(n,k,r)}{(n-2r)!} \frac{n!(2n-4r)!}{(n-2r)!(2n-2r)!}. \tag{3.84}$$

It clearly holds that

$$\frac{n!(2n-4r)!}{(n-2r)!(2n-2r)!} \leq 1, \tag{3.85}$$

hence

$$\begin{aligned} (3.84) &\leq \sum_{k=2r+1}^{n-2} \frac{f(n,k,r)}{(n-2r)!} \\ &= \left(\sum_{k=2r+1}^{2r+7} + \sum_{k=2r+8}^{\mathfrak{n}_\mathfrak{c}} + \sum_{k=\mathfrak{n}_\mathfrak{c}+1}^{n-2}\right) \frac{f(n,k,r)}{(n-2r)!} \\ &=: (A) + (B) + (C). \end{aligned} \tag{3.86}$$

By Proposition 10, and worst-case estimates, the following upperbounds hold:

$$
\begin{aligned}
(A) &\leq \sum_{k=2r+1}^{2r+7} \frac{(k+1)(n-k)!}{(n-2r)!} \leq \kappa_a \frac{7(2r+8)(n-2r-1)!}{(n-2r)!} \\
(B) &\leq \sum_{k=2r+8}^{\mathfrak{n}_\mathfrak{c}} \frac{n^6(n-k)!}{(n-2r)!} \leq n^6 \sum_{k=2r+8}^{\mathfrak{n}_\mathfrak{c}} \frac{(n-k)!}{(n-2r)!} \leq n^7 \frac{(n-2r-8)!}{(n-2r)!} \\
(C) &\leq \sum_{k=\mathfrak{n}_\mathfrak{c}+1}^{n-2} \frac{\left(2n^{7/8}\right)^{n-k}(n-k+1)}{(n-2r)!} \leq \frac{n^2\left(2n^{7/8}\right)^{5e(n+3)^{2/3}}}{(n-2r)!}\,.
\end{aligned}
\tag{3.87}
$$

All three terms are clearly vanishing in the limit $n \to \infty$. This implies that the second sum in (3.74) yields no contribution, and the proof of Lemma 9 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(λ)-distributed, i.e.

$$
\mathbb{P}(\widehat{W} = k|\mathcal{F})(\omega) = \frac{\lambda(\omega)^k}{k!} e^{-\lambda(\omega)}. \tag{3.88}
$$

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

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

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

$$
g_f(0) \equiv 0, \quad g_f(n) \equiv \frac{(n-1)!}{\lambda^n} \sum_{k=0}^{n-1} \frac{f(k)\lambda^k}{k!} \quad n > 0\,. \tag{3.89}
$$

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

$$
f(n) = \lambda g_f(n+1) - n g_f(n), \qquad n \geq 0\,, \tag{3.90}
$$

and

$$
g_f(n) = -\frac{(n-1)!}{\lambda^n} \sum_{k=n}^{\infty} \frac{f(k)\lambda^k}{k!} \qquad n > 0. \tag{3.91}
$$

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,

$$| g_f(n) |\leq \frac{(n-1)! \max_{k\in\mathbb{N}} | f(k) |}{\lambda^n} \int_0^\lambda \frac{t^{n-1}}{(n-1)!} e^t dt \leq \frac{\max_{k\in\mathbb{N}} | f(k) | e^\lambda}{n}. \tag{3.92}$$

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

$$f_{A,\lambda}(n) \equiv \mathbb{1}_{n\in A} - \mathbb{P}(\widehat{W} \in A|\mathcal{F}), \quad n \in \mathbb{N}. \tag{3.93}$$

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 (3.90), there exists a bounded $\mathcal{F}$-measurable function, denoted by $g_{A,\lambda}$, which satisfies

$$\mathbb{1}_{n\in A} - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) = \lambda g_{A,\lambda}(n+1) - n g_{A,\lambda}(n), \tag{3.94}$$

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

$$\mathbb{1}_{W\in A} - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) = \lambda g_{A,\lambda}(W+1) - W g_{A,\lambda}(W). \tag{3.95}$$

Taking conditional expectations thus yields

$$\begin{aligned} \mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) &= \lambda\mathbb{E}(g_{A,\lambda}(W+1)|\mathcal{F}) - \mathbb{E}(W g_{A,\lambda}(W)|\mathcal{F}) \\ &= \sum_{i\in I} \mathbb{E}(X_i|\mathcal{F})\mathbb{E}(g_{A,\lambda}(W+1)|\mathcal{F}) - \mathbb{E}(X_i g_{A,\lambda}(W)|\mathcal{F}). \end{aligned} \tag{3.96}$$

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.

$$\mathbb{P}(S^{(i)} = k|\mathcal{F}) = \mathbb{P}\left(\sum_{j\in N_i} X_j = k, X_i = 1 \Big| \mathcal{F}\right) \Big/ \mathbb{P}(X_i = 1 \big| \mathcal{F}). \tag{3.97}$$

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

$$\mathbb{E}(X_i g_{A,\lambda}(W)|\mathcal{F}) = \mathbb{P}(X_i = 1|\mathcal{F})\mathbb{E}\left[g_{A,\lambda}\left(1 + S^{(i)} + \sum_{j\in I\setminus(N_i\cup\{i\})} X_j\right) \Bigg| \mathcal{F}\right], \tag{3.98}$$

since X_i and X_j are conditionally independent given $\mathcal{F}$. Plugging this into the right-hand side of (3.96) yields

$$\begin{aligned}&\mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) \\ &\quad= \sum_{i\in I} \mathbb{E}(X_i|\mathcal{F})\mathbb{E}\left[g_{A,\lambda}(1+W) - g_{A,\lambda}(1+S^{(i)} + \sum_{j\in I\setminus(N_i\cup\{i\})} X_j) \middle| \mathcal{F}\right].\end{aligned} \tag{3.99}$$

Set now

$$M \equiv \sup\{|g_{A,\lambda}(n+1) - g_{A,\lambda}(n)| : n \in \mathbb{N}_0\}. \tag{3.100}$$

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

$$\begin{aligned}| \, \mathbb{P}(W \in A|\mathcal{F}) - \mathbb{P}(\widehat{W} \in A|\mathcal{F}) \, | &\leq M \sum_{i\in I} \mathbb{E}(X_i|\mathcal{F})\mathbb{E}(X_i + S^{(i)} + \sum_{j\in N_i} X_j|\mathcal{F}) \\ &= M\sum_{i\in I}\left(E(X_i|\mathcal{F})^2 + \sum_{j\in N_i} (\mathbb{E}(X_jX_i|\mathcal{F}) + \mathbb{E}(X_j|\mathcal{F})E(X_i|\mathcal{F}))\right).\end{aligned} \tag{3.101}$$

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

$$g_{A,\lambda} = \sum_{j\in A} g_{\{j\},\lambda}. \tag{3.102}$$

Furthermore,

$$\sum_{j=0}^{\infty} g_{\{j\},\lambda}(n+1) - g_{\{j\},\lambda}(n) = 0, \tag{3.103}$$

since by (3.102) it holds

$$\sum_{j=0}^{\infty} g_{\{j\},\lambda}(n) = g_{\mathbb{N}_0,\lambda}(n) = 0 \quad \forall n \in \mathbb{N}. \tag{3.104}$$

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

$$|g_{A,\lambda}(n+1) - g_{A,\lambda}(n)| \leq \sum_{j=0}^{\infty} (g_{\{j\},\lambda}(n+1) - g_{\{j\},\lambda}(n))^+. \tag{3.105}$$

By (3.89), the definition of f and elementary computations we have, for $0 < n \leq j$, that

$$g_{\{j\},\lambda}(n) = -\mathbb{P}(\widehat{W} = j|\mathcal{F}) \sum_{l=0}^{n-1} \frac{(n-1)!}{\lambda^{l+1}(n-1-l)!}. \tag{3.106}$$

This implies in particular that $g_{\{j\},\lambda}(n)$ is decreasing in n on $[0, j]$, hence all summands $j \geq n+1$ in (3.105) vanish. On the other hand, by (3.91), again the definition of f and elementary computations we have for $n > j$

$$g_{\{j\},\lambda}(n) = \mathbb{P}(\widehat{W} = j|\mathcal{F}) \sum_{l=0}^{\infty} \frac{\lambda^l (n-1)!}{(n+l)!}. \tag{3.107}$$

Since this is also decreasing in n, it follows that $j = n$ is the only non-zero summand in (3.105). All in all,

$$M = \sup_{n\in\mathbb{N}} \mid g_{A,\lambda}(n+1) - g_{A,\lambda}(n) \mid \leq \sup_{n\in\mathbb{N}} \mid g_{\{n\},\lambda}(n+1) - g_{\{n\},\lambda}(n) \mid . \tag{3.108}$$

Now, for $n > 0$, by (3.106) and (3.107),

$$\begin{aligned} &\mid g_{\{n\},\lambda}(n+1) - g_{\{n\},\lambda}(n) \mid = \\ &\quad = \frac{\lambda^n e^{-\lambda}}{n!} \left(\sum_{l=0}^{\infty} \frac{\lambda^l (n-1)!}{(n+l)!} + \sum_{l=0}^{n-1} \frac{(n-1)!}{\lambda^{l+1}(n-1-l)!} \right) \\ &\quad = \frac{e^{-\lambda}}{n} \left(\sum_{l=n}^{\infty} \frac{\lambda^l}{l!} + \sum_{l=0}^{n-1} \frac{\lambda^l}{l!} \right) = \frac{1}{n} \leq 1. \end{aligned} \tag{3.109}$$

On the other hand, for $n = 0$, we have

$$\mid g_{\{0\},\lambda}(1) - g_{\{0\},\lambda}(0) \mid = \frac{1}{\lambda}(1 - e^{-\lambda}) \leq 1 \tag{3.110}$$

by Taylor estimate. Using (3.109) and (3.110) in (3.108) shows that $M \leq 1$ as claimed, and concludes the proof of the conditional Chen-Stein method.

□

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

References.

[1] 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] D Barbour, Andrew, Lars Holst and Svante Janson. *Poisson approximation*, Oxford : Clarendon Press, (1992).

[3] Bolthausen, Erwin, and Nicola Kistler. *On a nonhierarchical version of the generalized random energy model, II: ultrametricity.* Stochastic Processes and their Applications 119.7 (2009): 2357-2386.

[4] Bovier, Anton. *Gaussian processes on trees: From spin glasses to branching Brownian motion.* Cambridge Studies in Advanced Mathematics Vol. 163, Cambridge University Press (2016).

[5] Bovier, Anton and Lisa Hartung. *The extremal process of two-speed branching Brownian motion.* Elect. J. Probab. 19, No. 18 (2014): 1-28.

[6] 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.

[7] Bovier, Anton and and Irina Kurkova. *A short course on mean field spin glasses.* In: A. Boutet de Monvel and A. Bovier (Eds.) Spin Glasses: Statics and Dynamics. Summer School Paris, 2007. Birkhäuser, Basel-Boston-Berlin, 2009

[8] Chiarini, Alberto, Alessandra Cipriani, and Rajat Hazra. *A note on the extremal process of the supercritical Gaussian Free Field*, Electron. Commun. Probab. 20 (2015), no. 74, 110.

[9] Derrida , Bernard and Herbert Spohn. *Polymers on disordered trees, spin glasses, and traveling waves.* J. Statist. Phys. 51, no. 5-6 (1988): 817840.

[10] Fang, Ming and Ofer Zeitouni. *Slowdown for time inhomogeneous branching Brownian motion.* J. Stat. Phys. 149 no. 1 (2012): 19.

[11] Fang, Ming and Ofer Zeitouni. *Branching random walks in time inhomogeneous environments.* Electron. J. Probab. 17 no. 67 (2012): 1-18.

[12] 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.

[13] 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).

[14] Kallenberg, Olav. *Random Measures, Theory and Applications.* Springer (2017).

[15] Kistler, Nicola, Adrien Schertzer and Marius A. Schmidt. *First passage percolation in the mean field limit.* Brazilian Journal of Probability and Statistics, to appear (2019+).

[16] Kistler, Nicola, Marius A. Schmidt. *From Derrida's random energy model to branching random walks: from 1 to 3.*. Elec. Comm. Prob. 20, no. 47, pp. 1-12 (2015).

[17] Martinsson, Anders. *First-passage percolation on Cartesian power graphs*, The Annals of Probability 2018, Vol. 46, No. 2, 10041041

[18] Villani, Cédric, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin (2009).

Institute for Mathematics, Goethe University Frankfurt a. M.
E-mail: kistler@math.uni-frankfurt.de
schertzer@math.uni-frankfurt.de

University of Basel, Switzerland
E-mail: mariusalexander.schmidt@unibas.ch