Adv. Appl. Prob. **48,** 1061–1094 (2016)
*doi:*10.1017/apr.2016.65

LARGE DEVIATION PRINCIPLES FOR CONNECTABLE RECEIVERS IN WIRELESS NETWORKS

CHRISTIAN HIRSCH,* **
BENEDIKT JAHNEL,*
PAUL KEELER * AND
ROBERT I. A. PATTERSON,* *Weierstrass Institute for Applied Analysis and Stochastics*

Abstract

We study large deviation principles for a model of wireless networks consisting of Poisson point processes of transmitters and receivers. To each transmitter we associate a family of connectable receivers whose signal-to-interference-and-noise ratio is larger than a certain connectivity threshold. First, we show a large deviation principle for the empirical measure of connectable receivers associated with transmitters in large boxes. Second, making use of the observation that the receivers connectable to the origin form a Cox point process, we derive a large deviation principle for the rescaled process of these receivers as the connection threshold tends to 0. Finally, we show how these results can be used to develop importance sampling algorithms that substantially reduce the variance for the estimation of probabilities of certain rare events such as users being unable to connect.

Keywords: Wireless network; signal-to-interference-and-noise ratio; large deviation principle; importance sampling

2010 Mathematics Subject Classification: Primary 60F10
Secondary 60K35

1. Model description and main results

We consider a stochastic geometry model for a wireless network consisting of a family of transmitters and a family of receivers. Transmitters and receivers are modeled by independent homogeneous Poisson point processes X and Y in $\mathbb{R}^d$ whose intensities are assumed to be nonzero and finite and will be denoted by λ_T and λ_R, respectively. For instance, we may think of transmitters and receivers as users participating in a device-to-device communication where messages need not be routed via a base station. It is believed that this form of communication will be a central concept in next-generation wireless networks [7]. The most basic requirement in the design of such networks is to guarantee satisfactory quality of service on average. Additionally, it is desirable to control and quantify the probability of a low quality of service occurring. This necessitates a more detailed probabilistic analysis and the theory of large deviations provides the appropriate tools.

Let us now describe the communication model. In order to determine the connection quality of messages sent out from a transmitter located at $x \in \mathbb{R}^d$ to a receiver located at $y \in \mathbb{R}^d$, the *signal-to-interference-and-noise ratio* (SINR) has been identified to be of fundamental importance [3]. More precisely, we assume that signals are transmitted with some positive powers P_x and decay according to the path-loss function $\ell(|x - y|)$, where '$|\cdot|$' denotes the

Received 4 June 2015; revision received 3 November 2015.
* Postal address: Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany.
** Email address: christian.hirsch@wias-berlin.de

Euclidean norm in $\mathbb{R}^d$ and $\ell\colon [0,\infty) \to [0,\infty)$ is a decreasing function satisfying $\ell(r) \in o(r^{-\alpha})$ for some $\alpha > d$. In particular, ℓ is bounded and $\int \ell(|x|)\,\mathrm{d}x < \infty$. In addition to the deterministic decay over distance, the signal strength is also influenced by random fading effects that are encoded in a positive random variable $F_{x,y}$. Such fading effects can, for example, arise from large obstacles in the environment or multipath interference due to moving reflectors [4, Chapter 22].

Furthermore, considering a signal sent out from X_i, the strength of the *interference* experienced at a location $y \in \mathbb{R}^d$ is assumed to be of the form

$$I(X_i, y) = I(X_i, y, X) = w + \sum_{j \neq i} P_{X_j} F_{X_j,y}\ell(|X_j - y|).$$

In words, the interference strength at a given location y consists of a contribution from the thermal noise $w > 0$ and the aggregated signal strengths coming from all other transmitters. For notational convenience, we differ from the common convention [3] and include the thermal noise w in the interference term. Hence, the SINR for the transmitter $X_i \in X$ and the possible receiver location $y \in \mathbb{R}^d$ is defined as the ratio of the signal strength by the interference, i.e.

$$\mathrm{SINR}(X_i, y) = \mathrm{SINR}(X_i, y, X) = \frac{P_{X_i} F_{X_i,y}\ell(|X_i - y|)}{I(X_i, y)}.$$

We assume that a connection can be established between $X_i \in X$ and $Y_j \in Y$ if $\mathrm{SINR}(X_i, Y_j) \geq t$ for some fixed connectivity threshold t. The importance of the SINR stems from Shannon's law in information theory, which provides an explicit formula expressing the maximum possible data throughput in terms of SINR, see [4, Chapter 16].

In this paper, we analyze how connectivity properties of the SINR-based network model described above behave in certain asymptotic regimes. First, we associate to each transmitter X_i the family of *receivers* $Y^{(i)}$ *that are connectable to* X_i, i.e.

$$Y^{(i)} = \{Y_j \in Y \colon \mathrm{SINR}(X_i, Y_j) \geq t\}.$$

An illustration of the transmitters together with their connectable receivers is shown in Figure 1. The family $Y^{(i)}$ can be used to express a variety of *frustration events* for the transmitter X_i. For instance, $\{Y^{(i)} = \varnothing\}$ describes the frustration event that the transmitter X_i is isolated, in the sense that it fails to communicate with any of the receivers. Similarly, if $B_r(X_i)$ denotes the open Euclidean ball with radius r centered at X_i, then $Y^{(i)} \subset B_r(X_i)$ encodes the event that X_i can only communicate with receivers at distance at most r.

Before we state our first main result, let us introduce the precise assumptions on the transmission powers and fading variables. We assume that the transmission powers $\{P_x\}_{x\in\mathbb{R}^d}$ form an independent and identically distributed (i.i.d.) random field whose existence is guaranteed by Kolmogorov's extension theorem. Note that only the subset of powers $\{P_{X_i}\}_{i\geq 1}$ is relevant, but it is notationally convenient to work with the random field indexed by the full space $\mathbb{R}^d$. A similar remark holds for the random fading field $\{F_{x,y}\}_{x,y\in\mathbb{R}^d}$. It can reproduce two different kinds of fading effects. First, a contribution stemming from a suitable random environment such as slow fading, which is typically spatially correlated. Second, effects such as fast fading, that are idiosyncratic to the pair (x, y) and, therefore, do not exhibit spatial correlation. To be more precise for the first contribution, we assume that Z is a homogeneous Poisson point process with intensity $\lambda_{\mathrm{E}} > 0$ modeling the random environment. Moreover, we use an i.i.d.

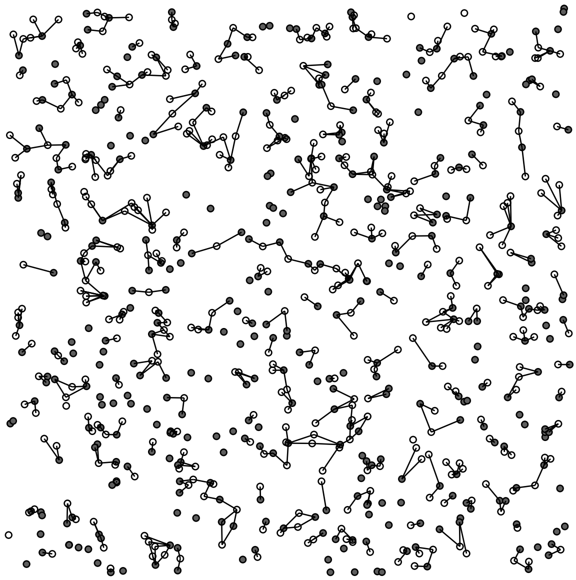

FIGURE 1: Realization of the network model. Transmitters (*solid circles*) are connected to receivers (*open circles*) by solid lines when a fixed SINR-threshold is exceeded.

random field $\{U_{x,y}\}_{x,y\in\mathbb{R}^d}$ consisting of random variables uniformly distributed on [0, 1] for the idiosyncratic effects. Then the random fading field can have the following general form:

$$F_{x,y} = \Phi(y-x, Z-x, U_{x,y}),$$

where Φ is measurable and positive. In particular, the construction is such that the fading field is spatially translation invariant, i.e. $\{F_{x+z,y+z}\}$ is equal in distribution to $\{F_{x,y}\}$ for any $z \in \mathbb{R}^d$.

The dependence of Φ on its second component should be local in the sense that there exists an increasing function $s_{\text{env}}\colon [0,\infty) \to [0,\infty)$ such that $\Phi(z,\varphi,u) = \Phi(z, \varphi \cap B_{s_{\text{env}}(|z|)}(o), u)$, where $B_{s_{\text{env}}(|z|)}(o)$ denotes the Euclidean ball of radius $s_{\text{env}}(|z|)$ centered at the origin. Moreover, letting U be a single uniformly distributed random variable on [0, 1], we assume that there exist $N > 0$, $s_{\max} > s_{\min} > 0$ such that for any $z \in \mathbb{R}^d$ and any locally finite $\varphi \subset \mathbb{R}^d$ the distribution function $q_{z,\varphi}\colon t \mapsto \mathbb{P}(1/\Phi(z,\varphi,U) \le t)$

- is globally Lipschitz with Lipschitz constant N,
- $q_{x,\varphi}(s) = 0$ for $s \le s_{\min}$ and $q_{x,\varphi}(s) = 1$ for $s > s_{\max}$.

The second condition ensures that the fading variables have support bounded away from 0 and ∞. We assume the same for the power variables P_x. Moreover the random objects X, Y, Z, $\{P_x\}$, and $\{U_{x,y}\}$ are independent.

We provide an example illustrating possible fading fields within the above framework. For instance, a Boolean model $\Xi = \bigcup_{Z_i \in Z} B_1(Z_i)$ can be interpreted as randomly distributed obstacles in a city. If the line of sight between transmitter x and receiver y is blocked by some building the signal propagation is diminished. That is, $F_{x,y} = \exp(-\mathbf{1}_{\{[x,y]\cap\Xi\neq\varnothing\}})J^{-1}(U_{\mathrm{x,y}})$, where J^{-1} is the generalized inverse of a globally Lipschitz function J that is the distribution function of a random variable which is bounded away from 0 and ∞. We note that for modeling

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

urban environments it is important to take into account the effects of correlated fading variables due to fixed obstacles [5].

Our first main result will provide a *large deviation principle* (LDP) for the empirical measure of the family of all connectable receivers $Y^{(i)} - X_i$ such that X_i is contained in the box $\Lambda_n = [-n/2, n/2]^d$ for large n. To make this precise, we first note that each $Y^{(i)}$ is a random variable in the measurable space $(\boldsymbol{N}_\mathrm{f}, \mathcal{N}_\mathrm{f})$. Here, $\boldsymbol{N}_\mathrm{f}$ is the family of all finite subsets of $\mathbb{R}^d$ that is endowed with the σ-algebra $\mathcal{N}_\mathrm{f}$ generated by maps of the form $\mathrm{ev}_B\colon \varphi \mapsto \#(\varphi \cap B)$ for any Borel set $B \subset \mathbb{R}^d$. In fact, $\boldsymbol{N}_\mathrm{f}$ is also a Polish space; see [9, Section A2.5]. Now, knowing the distribution of the *empirical measure*

$$L_n = \frac{1}{|\Lambda_n|} \sum_{X_i \in \Lambda_n} \delta_{Y^{(i)} - X_i},$$

we can answer questions such as:

- What is the probability that, when spatially averaged, a certain proportion of transmitters in Λ_n are isolated?
- What is the probability that, when spatially averaged, a certain proportion of transmitters in Λ_n have l receiver in an r proximity?

Apart from these examples, L_n can be used to describe more general events such as the average number of connectable receivers per transmitter, i.e. $|\Lambda_n|^{-1} \sum_{X_i \in \Lambda_n} \# Y^{(i)}$.

The empirical measure L_n is a random variable with values in the measurable space $(\mathcal{M}_\mathrm{f}(\boldsymbol{N}_\mathrm{f}), \mathcal{B}^{\mathrm{cy}}(\mathcal{M}_\mathrm{f}))$. Here, $\mathcal{M}_\mathrm{f}(\boldsymbol{N}_\mathrm{f})$ denotes the family of all finite measures on $\boldsymbol{N}_\mathrm{f}$ and $\mathcal{B}^{\mathrm{cy}}(\mathcal{M}_\mathrm{f})$ is the σ-algebra generated by the evaluation maps $\mu \mapsto \mu(B)$, where B is any bounded Borel set of $\boldsymbol{N}_\mathrm{f}$. Since our first main result provides a level-2 LDP, the τ-topology on $\mathcal{M}_\mathrm{f}$ will play an important role. This topology is generated by the maps $\mu \mapsto \mu(B)$, where B is any bounded Borel set of $\boldsymbol{N}_\mathrm{f}$. We refer the reader to [10, Section 6.2] for a detailed discussion of this topological space.

The LDP allows us to quantify the decay of probability for events away from their ergodic limit on an exponential scale. The exponential rate of decay to 0 is proportional to the volume and the proportionality factor is called the rate function. In order to identify the LDP rate function, we first recall the notion of *specific entropy of point marked random fields*. We follow the presentation in [16] and also refer the reader to [15, Chapter 15] for further details. Let E be a Polish space and write $\mathcal{E}$ for the corresponding Borel σ-algebra. Furthermore, let $\boldsymbol{N}_E$ denote the family of all configurations $\varphi \subset \mathbb{R}^d \times E$ whose projection to $\mathbb{R}^d$ is injective and with image forming a locally finite set. The space $\boldsymbol{N}_E$ is endowed with the smallest σ-algebra for which all evaluation maps $\varphi \mapsto \#(\varphi \cap (B \times F))$ are measurable for any Borel sets B, F of $\mathbb{R}^d$ and E, respectively. Any probability measure on $(\boldsymbol{N}_E, \mathcal{N}_E)$ is called an E-marked point random field. Let $n \geq 1$ and let P be an E-marked point random field whose realizations are contained in Λ_n with probability 1. Moreover, let Q be another E-marked point random field that is absolutely continuous with respect to P, where f denotes the respective density. Then, the *specific entropy* $H(Q \mid P)$ of Q with respect to P is defined as

$$H(Q \mid P) := P(f \log f),$$

where $P(f \log f)$ denotes the expectation of $f \log f$ with respect to P. This definition is extended to random point fields Q that are not absolutely continuous with respect to P by putting $H(Q \mid P) = \infty$. Finally, if P and Q are any E-marked point random fields, we introduce the notation

$$h(Q \mid P) := \sup_{n \geq 1} \frac{1}{|\Lambda_n|} H(Q_{\Lambda_n} \mid P_{\Lambda_n}),$$

where P_{Λ_n} and Q_{Λ_n} denote the projection of P and Q to Λ_n.

In the following, we write $\mathcal{P}_\theta$ for the family of all stationary E-marked point random fields of finite intensity. Here, a stationary E-marked point random field is a probability measure on $\boldsymbol{N}_E$ that is invariant with respect to shifts on $\mathbb{R}^d$. The intensity of Q is defined as

$$\int_{\boldsymbol{N}_E} \#\{(x_i, e_i) \in \varphi\colon\ x_i \in [0, 1]^d\} Q(\mathrm{d}\varphi).$$

We also need the notion of the Palm version of a stationary point random field as defined, for example, in [20]. The (unnormalized) *Palm mark measure* Q^o associated with $Q \in \mathcal{P}_\theta$ is given by

$$Q^o(F) = \int_{\boldsymbol{N}_E} \#\{(x_i, e_i) \in \varphi\colon\ (x_i, e_i) \in [0, 1]^d \times F\} Q(\mathrm{d}\varphi), \qquad F \in \mathcal{E}.$$

In other words, after normalization, Q^o describes the distribution of the marks of Q.

The concept of random marked point random fields is very flexible so that the probability space associated with X, Y, Z, $\{P_x\}$, and $\{U_{x,y}\}$ can be encoded in this framework, see Section 2.2 for details.

Let us state the first main result of this paper, an LDP for the empirical measure of connectable receivers associated with transmitters in a large box. Starting from a stationary point random field $\mathbb{Q}$ of transmitters, receivers, and environment, we define $\mathbb{Q}^*$ as the Palm mark measure of the stationary $\boldsymbol{N}_\mathrm{f}$-marked point random field defined by $\{(X_i, Y^{(i)} - X_i)\}_{i \geq 1}$.

Theorem 1. *The random measures $\{L_n\}_{n \geq 1}$ satisfy an LDP in the τ-topology with rate $|\Lambda_n|$ and good rate function*

$$\mathcal{I}(Q) = \inf_{\mathbb{Q} \in \mathcal{P}_\theta,\, \mathbb{Q}^* = Q} h(\mathbb{Q} \mid \mathbb{P}).$$

That is, for all $A \in \mathcal{B}^{\mathrm{cy}}(\mathcal{M}_\mathrm{f})$,

$$\limsup_{n \to \infty} \frac{1}{|\Lambda_n|} \log \mathbb{P}(L_n \in A) \leq - \inf_{Q \in \bar{A}} \mathcal{I}(Q)$$

and

$$\liminf_{n \to \infty} \frac{1}{|\Lambda_n|} \log \mathbb{P}(L_n \in A) \geq - \inf_{Q \in A^o} \mathcal{I}(Q),$$

where $\bar{A}$ denotes the closure and A^o the interior of A. Moreover, the function $\mathcal{I}$ is lower semi-continuous and has compact level sets.

To prove Theorem 1, we make use of the level-3 LDPs established in [16] (see also [13] and [14] for related results). However, the long-range dependencies induced by the interferences prevent us from applying the contraction principle directly. Similarly to [1], we first have to perform a truncation step and consider an approximate model with finite-range dependencies. In order to deduce Theorem 1 from the level-3 LDP in the truncated scenario, we show that by a suitable choice of the truncation range, the truncation error becomes arbitrarily small. We note that a certain finite-range approximation has also been used in the proof of LDPs for stabilizing functionals [22]. Although functionals involving SINR do not fit into the framework of stabilization, an alternative approach to proving Theorem 1 could be to try and extend the arguments in [22], which are based on the notion of near-additivity [23], to the present setting.

In our second main result, we investigate how the connectable receivers associated with a typical transmitter located at the origin behave as the connection threshold t tends to 0. Since this scenario turns out to be more complicated than the one considered in Theorem 1, we impose stronger additional assumptions. To be more precise, we assume that $\ell(r) = \min\{1, r^{-\alpha}\}$ for some $\alpha > d$, the transmission power at the origin is fixed (say equal to 1) and that there is no random environment Z. That is, $\{F_{x,y}\}_{x,y\in\mathbb{R}^d}$ are i.i.d. and we set $q(a) = \mathbb{P}(1/F_{x,y} \le a)$. Moreover, we assume that there exist $N > 0$ and $s_{\min} > 0$ such that

- q is globally Lipschitz and globally Lipschitz in its first derivative, both with Lipschitz constant N,
- $q(s) = 0$ for $s \le s_{\min}$ and $q(s) > 0$ for $s > s_{\min}$.

To begin with, we provide some important preliminary observations: first, we note that the receivers connectable to the origin; namely,

$$Y^t = \{Y_j \in Y \colon \mathrm{SINR}(o, Y_j, X \cup \{o\}) \ge t\},$$

form a Cox point process with random intensity measure M_t given by

$$M_t(B) = \lambda_{\mathrm{R}} \int_B \Gamma(t^{-1}\ell(|y|), y)\,\mathrm{d}y,$$

where

$$\Gamma(a, y) = \mathbb{E}(q(aI(y)^{-1}) \mid X) \quad \text{for } a \ge 0 \text{ and } y \in \mathbb{R}^d. \tag{1}$$

In other words, Γ is an expectation with respect to the fading field in the interference. More precisely, it is the conditional expectation on the transmitter process $X = \{(X_i, P_i)\}$ that also carries the transmission powers as marks. For instance, this observation implies that the probability for the origin to be isolated is given by $p_t = \mathbb{E}\exp(-M_t(\mathbb{R}^d))$ and tends to 0 as t tends to 0. The representation of the isolation probability provides a strong hint that the Varadhan–Laplace technique from the theory of large deviations (see, e.g. [16]) could be a useful tool in the analysis of the asymptotic behavior of p_t as t tends to 0. In particular, p_t should decay exponentially as t tends to 0. The exact form of this decay is presented in Corollary 1. In Theorem 2, we give a more general result describing the exponential decay of unlikely numbers of connectable receivers in space. Throughout the entire manuscript, $\beta = 1/\alpha$ denotes the inverse of the path-loss exponent. Furthermore, we set $\Lambda'_t = \Lambda_{2(ws_{\min}t)^{-\beta}}$, so that $q(t^{-1}\ell(|y|)I(y)^{-1}) = 0$ if $y \notin \Lambda'_t$ and write 'Po_s' for the stationary point random field induced by a homogeneous Poisson point process with intensity $s \ge 0$.

Theorem 2. *The random measures* $\{|\Lambda'_t|^{-1} Y^t(t^{-\beta}\cdot)\}_{t<1}$ *satisfy an LDP in the weak topology with rate* $|\Lambda'_t|$ *and good rate function given by*

$$\mathcal{I}(\varphi) = \begin{cases} \displaystyle\int_{\Lambda'_1} \mathcal{I}_y(\dot\varphi(y))\,\mathrm{d}y & \text{if } \mathrm{d}\varphi/\mathrm{d}x = \dot\varphi \text{ exists,} \\ \infty & \text{otherwise,} \end{cases}$$

where

$$\mathcal{I}_y(s) = \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + h(\mathrm{Po}_s \mid \mathrm{Po}_{\lambda_\mathrm{R}\mathbb{Q}(\Gamma(|y|^{-\alpha},o))})), \tag{2}$$

and $\mathbb{Q}(\Gamma(|y|^{-\alpha}, o))$ *denotes the expectation of* $\Gamma(|y|^{-\alpha}, o)$ *for a stationary marked point process* $X = \{(X_i, P_i)\}$ *that is distributed according to* $\mathbb{Q}$*.*

In contrast to Theorem 1, the probability measures $\mathbb{Q} \in \mathcal{P}_\theta$ in (2) are distributions only of the transmitters X and their transmission powers P. Setting $\varphi \equiv 0$ gives the decay of isolation probability, this is the content of the following corollary.

Corollary 1. *We have*

$$\lim_{t\to 0} |\Lambda'_t|^{-1} \log p_t = \lim_{t\to 0} |\Lambda'_t|^{-1} \log \mathbb{E} \exp\bigg(-\lambda_\mathrm{R} \int_{\mathbb{R}^d} \Gamma(t^{-1}\ell(|y|), y)\, \mathrm{d}y\bigg)$$
$$= -\int_{\Lambda'_1} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_\mathrm{R}\mathbb{Q}(\Gamma(|y|^{-\alpha}, o)))\, \mathrm{d}y.$$

Large deviation principles in SINR-based networks have already been considered in [12] and [25]. However, the question treated in Theorem 2 is in a certain sense dual to the ones discussed in [12] and [25]. In those papers a large deviation principle was derived for the interference at the origin caused by the signals from other users. We investigate a scenario where the origin sends out a signal and we are interested in the interference at the location of the other users.

To prove Theorem 2, we introduce a stationary point process that carries more information than Y^t. For this point process, we first establish a level-1 LDP based on the results of [16], and then deduce a path-space LDP using the Dawson–Gärtner technique. The proof is concluded by applying the contraction principle.

In Corollary 1 we showed that p_t decays exponentially in $t^{-d/\alpha}$ and provides a variational characterization of the rate function. However, for the purpose of estimating the actual value of p_t, our asymptotic result has two drawbacks. First, in Corollary 1, we do not make any claims as regards to how small t should be for the asymptotic to be an acceptable approximation. It is not at all clear from the variational formula how to compute (or even approximate) the asymptotic rate function. Nevertheless, when estimating the isolation probability p_t via Monte Carlo simulations, our large deviation result can be used to devise an importance sampling scheme that substantially reduces the estimation variance. In the field of stochastic processes, large deviation techniques have emerged as a powerful tool to find suitable importance sampling densities [2, Chapter 6.6], but so far have not found widespread use for spatial rare-event problems.

As a notable exception, we mention [24], which deals with rare events arising from large values of the interference measured at the origin. In that paper, it was shown that the asymptotically efficient importance sampling density is given by a certain inhomogeneous Poisson point process. In our setting, the variational characterization in Theorem 2 suggests that the asymptotically optimal density is not given by a Poisson point process, but by a collection of location-dependent Gibbs processes. Still, in a first step, we provide simulation results illustrating that using an isotropic Poisson point process already leads to substantial variance reduction. Let us also note that importance sampling for Gibbs processes on the lattice has been studied in [6].

The present paper is organized as follows. In Sections 2 and 3 we provide the proofs for Theorems 1 and 2, respectively. Section 3 also contains the proof of Corollary 1. Finally, in Section 4 we describe two importance sampling schemes and provide some simulation results.

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

2. Proof of Theorem 1

As mentioned in Section 1, in order to prove Theorem 1, we use the classical level-3 large deviation result for Poisson point processes [16, Theorem 3.1]. However, the interferences induce long-range interactions that are not immediately compatible with the topology $\tau_{\mathcal{L}}$ of local convergence that is used in [16]. To resolve this issue, we will proceed similarly to [1] and show that a suitable truncation of the path-loss functions appearing in the interference expression induces only a negligible error; see Section 2.1. After this truncation, we show in Section 2.2 how the LDP for the stationary empirical field [16, Theorem 3.1] can be used to prove Theorem 1.

2.1. Truncation of the path-loss function

First, we show that only an asymptotically negligible error occurs when disregarding transmitters close to the boundary of Λ_n. This is a well-known consequence of the Poisson concentration property [8, Chapter 2.2], but for the convenience of the reader, we provide a detailed proof.

Lemma 1. *Let $b, \varepsilon > 0$ be arbitrary. Then,*

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \mathbb{P}(X(\Lambda_n \setminus \Lambda_{n-b}) \geq \varepsilon|\Lambda_n|) = -\infty.$$

Proof. Let $\delta = \lambda_{\mathrm{T}}(1-(1-b/n)^d)$, $m = \delta n^d$, and $\tau = \varepsilon n^d$, then the Poisson concentration inequality [8, Chapter 2.2] implies that

$$\begin{aligned}\mathbb{P}(X(\Lambda_n \setminus \Lambda_{n-b})) \geq \tau) &\leq \left(\frac{m}{\tau}\right)^{\tau} \exp(\tau - n)\\ &= (\delta\varepsilon^{-1})^{\varepsilon n^d} \exp((\varepsilon - \delta)n^d)\\ &\leq \exp(\varepsilon n^d \log(e\delta\varepsilon^{-1})).\end{aligned}$$

Since $\log(e\delta\varepsilon^{-1})$ tends to $-\infty$ as $n \to \infty$, this proves the claim. □

Next, we show that truncating the path-loss function in the interference at a finite threshold only leads to a small error provided that the threshold is chosen sufficiently large. To be more precise, for $b \geq 1$ we set $\ell_b(r) = \ell(r)$ if $r < b$ and $\ell_b(r) = 0$ if $r \geq b$. Furthermore, we define

$$I_b(X_i, y) = w + \sum_{j\neq i} P_{X_j} F_{X_j,y} \ell_b(|X_j - y|), \qquad \mathrm{SINR}_b(X_i, y) = \frac{P_{X_i} F_{X_i,y} \ell(|X_i - y|)}{I_b(X_i, y)},$$

and

$$L_n^b = \frac{1}{|\Lambda_n|} \sum_{X_i\in\Lambda_n} \delta_{Y^{(i),b} - X_i},$$

where $Y^{(i),b} = \{Y_j \in Y : \mathrm{SINR}_b(X_i, Y_j) \geq t\}$ denotes the point process of *b-connectable receivers* for the transmitter X_i. We show that when using the total variation distance

$$d_{\mathrm{TV}}(L_n, L_n^b) = \sup_{B\in\mathcal{N}_f} |L_n(B) - L_n^b(B)|,$$

the random measures $\{L_n^b\}_{n\geq1}$ are exponentially good approximations of the random measures $\{L_n\}_{n\geq1}$ in the sense of [10, Definition 4.2.14].

Lemma 2. *Let $\varepsilon > 0$ be arbitrary. Then,*

$$\lim_{b\to\infty}\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}\log\mathbb{P}(d_{\mathrm{TV}}(L_n, L_n^b)\ge\varepsilon) = -\infty.$$

Proof. To be specific and for notational convenience, let us assume that the support of the power variables is contained in $[s_{\max}^{-1}, s_{\min}^{-1}]$. The definition of the total variation distance implies that

$$d_{\mathrm{TV}}(L_n, L_n^b)\le\frac{1}{|\Lambda_n|}\#\{X_i\in\Lambda_n\colon Y^{(i)}\ne Y^{(i),b}\}.$$

Next, by Lemma 1, we only need to consider those X_i that are contained in Λ_{n-2r_0}, where $r_0 > 0$ is chosen such that $\ell(r_0)\le ws_{\min}^2 t$. Then, almost surely, for $X_i\in\Lambda_{n-2r_0}$ and $Y_i\in\Lambda_n^c$, $\mathrm{SINR}_b(X_i, Y_j) < t$ for all $b\ge 1$. Consequently, it suffices to bound the number of transmitter-receiver pairs $(X_i, Y_j)\in X\times Y$ such that $X_i\in\Lambda_{n-2r_0}$, $Y_i\in\Lambda_n$, and $\mathrm{SINR}(X_i, Y_j) < t \le \mathrm{SINR}_b(X_i, Y_j)$. In fact, it suffices to focus on the receivers in these pairs. Indeed, let us call Y_j *b-pivotal* if there exists some transmitter X_i such that the pair (X_i, Y_j) has these properties. Then, since we assumed that $q_{x,\varphi}(r) = 0$ for $r\le s_{\min}$, for each receiver Y_j there exist $K = \lceil t^{-1}s_{\max}^2/s_{\min}^2\rceil$ transmitters, $A(Y_j, X) = \{X_{i_1},\dots,X_{i_K}\}$ such that $\mathrm{SINR}_b(X_i, Y_j) < t$ if $X_i\notin A(Y_j, X)$. Hence, it suffices to show that, for every $\varepsilon > 0$,

$$\lim_{b\to\infty}\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}\log\mathbb{P}(\#\{Y_j\in\Lambda_n\colon Y_j\text{ is }b\text{-pivotal}\}\ge\varepsilon|\Lambda_n|) = -\infty.$$

In order to do so, we use the exponential Markov inequality with $s\ge 1$ and estimate

$$\begin{aligned}&\mathbb{P}(\#\{Y_j\in\Lambda_n\colon Y_j\text{ is }b\text{-pivotal}\}\ge\varepsilon|\Lambda_n|)\\&\qquad\le\exp(-s\varepsilon|\Lambda_n|)\mathbb{E}\exp(s\#\{Y_j\in\Lambda_n\colon Y_j\text{ is }b\text{-pivotal}\}).\end{aligned}$$

Hence, it suffices to show that, for every $s\ge 1$,

$$\lim_{b\to\infty}\limsup_{n\to\infty}\frac{1}{|\Lambda_n|}\log\mathbb{E}\exp(s\#\{Y_j\in\Lambda_n\colon Y_j\text{ is }b\text{-pivotal}\}) = 0.$$

The point process of receivers that are b-pivotal form a stationary Cox point process with random intensity measure

$$M'(B) = \lambda_{\mathrm{R}}\int_B\mathbb{P}(y\text{ is }b\text{-pivotal}\mid X, Z)\,\mathrm{d}y,\qquad B\in\mathcal{B}(\mathbb{R}^d),$$

where we think of $X = \{(X_i, P_i)\}_{i\ge 1}$ as a marked point process and we evaluate the probability with respect to the fading variables associated with the pairs $(y, X_i)_{i\ge 1}$. Since q is assumed to be globally Lipschitz with constant N, we arrive at

$$\begin{aligned}&\mathbb{P}(y\text{ is }b\text{-pivotal}\mid X, Z)\\&\quad\le\sum_{X_i\in A(y,X)}\mathbb{P}(\mathrm{SINR}(X_i, y) < t\le\mathrm{SINR}_b(X_i, y)\mid X, Z)\\&\quad\le\sum_{X_i\in A(y,X)}\mathbb{P}(F_{X_i,y}^{-1}t\in[P_{X_i}\ell(|X_i-y|)I(X_i, y)^{-1},\\&\qquad\qquad\qquad P_{X_i}\ell(|X_i-y|)I_b(X_i, y)^{-1}]\mid X, Z)\\&\quad\le\sum_{X_i\in A(y,X)}\ell(|X_i-y|)Ns_{\min}^{-1}t^{-1}w^{-2}\mathbb{E}(I(X_i, y)-I_b(X_i, y)\mid X, Z),\end{aligned}$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

which is at most $S\sum_{i\geq 1}\ell(|X_i-y|)-\ell_b(|X_i-y|)$, where $S = K\ell(0)Ns_{\min}^{-3}t^{-1}w^{-2}$. In particular, we obtain

$$M'(B) \leq S' \int_B \sum_{i\geq 1} \ell(|X_i-y|)-\ell_b(|X_i-y|)\,\mathrm{d}y,$$

where $S' = \lambda_{\mathrm{R}} S$. Hence, using the formula for the Laplace functional of a Cox point process, we obtain

$$\begin{aligned}
&\mathbb{E}\exp[s\#\{Y_j\in\Lambda_{2n}\colon Y_j \text{ is } b\text{-pivotal}\}]\\
&\quad\leq \mathbb{E}\exp\bigg[(\mathrm{e}^s-1)S'\int_{\Lambda_{2n}}\sum_{i\geq 1}\ell(|X_i-y|)-\ell_b(|X_i-y|)\,\mathrm{d}y\bigg]\\
&\quad= \exp\bigg[\lambda_{\mathrm{T}}\int_{\mathbb{R}^d}\exp\bigg((\mathrm{e}^s-1)S'\int_{\Lambda_{2n}}\ell(|x-y|)-\ell_b(|x-y|)\,\mathrm{d}y\bigg)-1\,\mathrm{d}x\bigg].
\end{aligned}$$

Note, that we can bound the integral

$$\begin{aligned}
&\int_{\mathbb{R}^d}\exp\bigg[\tau\int_{\Lambda_{2n}}\ell(|x-y|)-\ell_b(|x-y|)\,\mathrm{d}y\bigg]-1\,\mathrm{d}x\\
&\qquad\leq \int_{\Lambda_{4n}}\exp\bigg[\tau\int_{\mathbb{R}^d\setminus B_b(x)}\ell(|x-y|)\,\mathrm{d}y\bigg]-1\,\mathrm{d}x\\
&\qquad\quad+\int_{\mathbb{R}^d\setminus\Lambda_{4n}}\exp\bigg[\tau\int_{\Lambda_{2n}}\ell(|x-y|)\,\mathrm{d}y\bigg]-1\,\mathrm{d}x,
\end{aligned}$$

where $\tau := (\mathrm{e}^s-1)S'$. In the next step, we derive bounds for these expressions separately. For the first, we have

$$\frac{1}{n^d}\int_{\Lambda_{4n}}\exp\bigg[\tau\int_{\mathbb{R}^d\setminus B_b(x)}\ell(|x-y|)\,\mathrm{d}y\bigg]-1\,\mathrm{d}x = 4^d\bigg[\exp\bigg(\tau\int_{\mathbb{R}^d\setminus B_b(o)}\ell(|y|)\,\mathrm{d}y\bigg)-1\bigg],$$

which tends to 0 as b tends to ∞. For the second expression, we note that, for $x\in\mathbb{R}^d\setminus\Lambda_{4n}$ and $y\in\Lambda_{2n}$,

$$|x-y| = \frac{|x-y|+\sqrt{d}|x-y|}{1+\sqrt{d}} \geq \frac{|x-y|+|y|}{1+\sqrt{d}} \geq (1+\sqrt{d}\,)^{-1}|x|.$$

Consequently, using the fact that $\ell(r)\in o(r^{-d})$, we have, for large n,

$$\begin{aligned}
&\frac{1}{n^d}\int_{\mathbb{R}^d\setminus\Lambda_{4n}}\exp\bigg[\tau\int_{\Lambda_{2n}}\ell(|x-y|)\,\mathrm{d}y\bigg]-1\,\mathrm{d}x\\
&\qquad\leq \frac{1}{n^d}\int_{\mathbb{R}^d\setminus\Lambda_{4n}}\exp[\tau(2n)^d\ell((1+\sqrt{d}\,)^{-1}|x|)]-1\,\mathrm{d}x\\
&\qquad\leq \int_{\mathbb{R}^d\setminus\Lambda_{4n}}\tau 2^{d+1}\ell((1+\sqrt{d}\,)^{-1}|x|)\,\mathrm{d}x\\
&\qquad= \int_{\mathbb{R}^d\setminus\Lambda_{4n(1+\sqrt{d}\,)^{-1}}}\tau 2^{d+1}(1+\sqrt{d}\,)^d\ell(|x|)\,\mathrm{d}x,
\end{aligned}$$

which tends to 0 as n tends to ∞. □

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

2.2. Application of an LDP for the stationary empirical field

In order to apply [16, Theorem 3.1], we need to relate the empirical measure of connectable receivers to the stationary empirical field considered in [16]. Here, the first task consists of encoding the probability space carrying the point processes of transmitters X, the point process of receivers Y, the random environment Z, the transmission powers $\{P_x\}$, and the i.i.d. family $\{U_{x,y}\}$ in the framework of stationary marked point processes. To be more precise, we set $\Sigma = \{\mathrm{E}, \mathrm{R}, \mathrm{T}\}$ and consider the mark space $E = \Sigma \times (0, \infty) \times [0, 1]^{\mathbb{N}}$ equipped with some complete and separable metric. Furthermore, we let V denote an independently E-marked homogeneous Poisson point process with intensity $\sum_{\sigma\in\Sigma} \lambda_\sigma$. The mark distribution on E is a product of three distributions defined on the spaces Σ, $(0, \infty)$, and $[0, 1]^{\mathbb{N}}$, respectively. First, on Σ, we choose the distribution assigning $\sigma \in \Sigma$ the probability $\lambda_\sigma / \sum_{\sigma'\in\Sigma} \lambda_{\sigma'}$. Second, on $(0, \infty)$, we choose the distribution of the transmission power P_x considered in Section 1. Third, the distribution on $[0, 1]^{\mathbb{N}}$ describes a family of i.i.d. random variables that are uniformly distributed $[0, 1]$. The Poisson point process Z that generates the random environment is represented by elements of $V = (v_i, \sigma_i, P_i, (U_{i,j})_{j\geq 1})_{i\geq 1}$ with $\sigma_i = \mathrm{E}$. Elements of $V = (v_i, \sigma_i, P_i, (U_{i,j})_{j\geq 1})_{i\geq 1}$ with $\sigma_i = \mathrm{T}$ are thought of as transmitters and are denoted by X. Elements of $V = (v_i, \sigma_i, P_i, (U_{i,j})_{j\geq 1})_{i\geq 1}$ with $\sigma_i = \mathrm{R}$ are thought of as receivers and are denoted by Y. We note that the power variables P_i have no meaning if $\sigma_i \neq \mathrm{T}$. The random variables $U_{x,y}$ should be thought of as being attached to the transmitters. Moreover, proceeding as in [16, Section 1], let

$$V^{\mathrm{per},n} = \bigcup_{s\in\mathbb{Z}^d} ((V \cap \Lambda_n) + ns)$$

denote the periodic spatial continuation of $V \cap \Lambda_n$. The *stationary empirical field* is defined as

$$R_{n,V} := \frac{1}{|\Lambda_n|} \int_{\Lambda_n} \mathbf{1}_{\{V^{\mathrm{per},n} - v\}}\, \mathrm{d}v,$$

where $V^{\mathrm{per},n} - v = \{(v_j - v, e_j)\}_{j\geq 1}$ is the spatial translation of $V^{\mathrm{per},n}$ by v. Now, we let $Y^{\mathrm{per},n,b,(i)}$ denote the family of periodized receivers that have a b-connection to the transmitter $X_i^{\mathrm{per},n} = (x_i, \mathrm{T}, P_i, (U_{i,l})_{l\geq 1})$. More precisely,

$$Y^{\mathrm{per},n,b,(i)} = \biggl\{ Y_j = (y_j, \mathrm{R}, P_j, (U_{j,l})_{l\geq 1}) \in Y^{\mathrm{per},n} : t \leq \frac{P_i F_{x_i,y_j} \ell(|x_i - y_j|)}{w + \sum_{k\neq i} P_k F_{x_k,y_j} \ell(|x_k - y_j|)} \biggr\},$$

where

$$F_{x_i,y_j} = \Phi(y_j - x_i, Z^{\mathrm{per},n} - x_i, U_{i,\Psi(y_j - x_i, Y^{\mathrm{per},n} - x_i)}),$$

and where the integer $\Psi(y_j - x_i, Y^{\mathrm{per},n} - x_i) \geq 1$ is defined as follows. If $k \geq 1$ is such that $y_j - x_i$ is the kth closest element in $Y^{\mathrm{per},n} - x_i$ to the origin, then we set $\Psi(y_j - x_i, Y^{\mathrm{per},n} - x_i) = k$. This construction will ensure translation invariance for the periodized version. The empirical measure $L_n^{\mathrm{per},b}$ of b-connectable receivers associated with transmitters in Λ_n when the network is based on periodized configurations can also be expressed as a function of $R_{n,V}$. Indeed, by the same technique that was used to define the individual empirical field in [16], we arrive at

$$L_n^{\mathrm{per},b} = \frac{1}{|\Lambda_n|} \sum_{x_i \in X^{\mathrm{per},n} \cap \Lambda_n} \delta_{Y^{\mathrm{per},n,b,(i)} - x_i} = \frac{1}{|\Lambda_n|} \int_{\Lambda_n} g'(V^{\mathrm{per},n} - v)\, \mathrm{d}v,$$

where

$$g'(V^{\mathrm{per},n} - v) = \sum_{x_i - v \in (X^{\mathrm{per},n} - v) \cap \Lambda_1} g(V^{\mathrm{per},n} - x_i),$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

and where g is the Dirac measure concentrated on the family of b-connectable receivers from the origin multiplied with the indicator function that the origin is a transmitter. Next, we prove that the random measures $\{L_n^{\mathrm{per},b}\}_{n\geq 1}$ and $\{L_n^b\}_{n\geq 1}$ are exponentially equivalent (in the sense of [10, Definition 4.2.10]), when using the total variation metric.

Lemma 3. *Let $\varepsilon > 0$ be arbitrary. Then,*

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \mathbb{P}(d_{\mathrm{TV}}(L_n^{\mathrm{per},b}, L_n^b) \geq \varepsilon) = -\infty.$$

Proof. As in Lemma 2, choose $r_0 \geq 1$ such that $\ell(|x-y|) \leq w s_{\min}^2 t$ if $|x-y| \geq r_0$. In particular, $Y^{(i),b} \subset B_{r_0}(X_i)$. Furthermore, by the truncation of the interference, to decide whether $Y_j \in Y^{(i),b}$ it suffices to look at transmitters in $B_b(Y_j)$. As a consequence, the family $Y^{(i),b}$ depends only on the network configuration in $B_{r_0+b+s_{\mathrm{env}}(b)}(X_i)$. Hence,

$$d_{\mathrm{TV}}(L_n^{\mathrm{per},b}, L_n^b) \leq \#(X \cap \Lambda_n \setminus \Lambda_{n-2(r_0+b+s_{\mathrm{env}}(b))}),$$

and the claim follows from Lemma 1. □

Now, we are in a position to provide a proof for the LDP asserted in Theorem 1 when L_n is replaced by L_n^b. Let $\mathbb{Q}$ be the distribution of some stationary E-marked point process $V = (v_i, \sigma_i, P_i, (U_{i,j})_{j\geq 1})_{i\geq 1}$. Then, we define $\mathbb{Q}^{*,b}$ as the Palm mark measure of the marked point process $(X_i, Y^{(i),b} - X_i)$. Here, as above, $X_i \in V$ are interpreted as transmitters and $Y^{(i),b} \subset V$ as the b-connectable receivers.

Proposition 1. *The random measures $\{L_n^b\}_{n\geq 1}$ satisfy an LDP in the τ-topology with rate $|\Lambda_n|$ and good rate function*

$$\mathcal{I}^b \colon Q \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta,\, \mathbb{Q}^{*,b}=Q} h(\mathbb{Q} \mid \mathbb{P}).$$

Proof. First, we note that the map $\Phi_b \colon \mathbb{Q} \mapsto \mathbb{Q}^{*,b}$ is continuous with respect to the τ-topology. Indeed, for $\mathbb{Q}_n \to \mathbb{Q}$ and $B \in \mathcal{N}_f$, the locality that is established after truncating the interferences leads to $|\mathbb{Q}^{*,b}(B) - \mathbb{Q}_n^{*,b}(B)| \to 0$ as $n \to \infty$. As $\Phi_b(R_{n,V}) = L_n^{\mathrm{per},b}$, we can apply [16, Corollary 3.2] and the contraction principle. Thus, the random measures $\{L_n^{\mathrm{per},b}\}_{n\geq 1}$ satisfy an LDP with good rate function $\mathcal{I}_b$. Finally, in Lemma 3 we show that $\{L_n^{\mathrm{per},b}\}_{n\geq 1}$ and $\{L_n^b\}_{n\geq 1}$ are exponentially equivalent with respect to the total variant distance. This implies the exponential equivalence of $\{L_n^{\mathrm{per},b}\}_{n\geq 1}$ and $\{L_n^b\}_{n\geq 1}$ when evaluated on an arbitrary Borel subset of $\boldsymbol{N}_{\mathrm{f}}$. So the claim follows from [11, Corollary 1.10, Remark 1.4]. □

The same arguments also prove the following result, where we consider the marked point process $(X_i, Y^{(i),b} - X_i, Y^{(i),b'} - X_i)$ at different truncation thresholds $b' > b \geq 1$. Starting from $\mathbb{Q} \in \mathcal{P}_\theta$, the associated Palm mark distribution is denoted by $\mathbb{Q}^{*,b,b'}$.

Lemma 4. *Let $b' > b \geq 1$. Then, the random variables $\{(1/|\Lambda_n|)\#\{X_i \in \Lambda_n \colon Y^{(i),b'} \neq Y^{(i),b}\}\}_{n\geq 1}$ satisfy an LDP with rate $|\Lambda_n|$ and good rate function*

$$s \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta,\, \mathbb{Q}^{*,b,b'}(Y^{(o),b'}\neq Y^{(o),b})=s} h(\mathbb{Q} \mid \mathbb{P}).$$

Finally, we complete the proof of Theorem 1. In Lemma 2, we showed that $\{L_n^b\}_{n\geq 1}$ are exponentially good approximations of $\{L_n\}_{n\geq 1}$ and, hence, an application of [11, Theorem 1.13] is natural.

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

Proof of Theorem 1. As mentioned in the previous paragraph, [11, Theorem 1.13] implies that it suffices to verify the following condition. For every $\varepsilon, K > 0$ there exists $b \geq 1$ such that $\sup_{\mathbb{Q}\in\mathcal{P}_\theta,\, h(\mathbb{Q}\mid\mathbb{P})\leq K} d_{\mathrm{TV}}(\mathbb{Q}^*, \mathbb{Q}^{*,b}) \leq \varepsilon$. We show that a slightly stronger statement holds, where $d_{\mathrm{TV}}(\mathbb{Q}^*, \mathbb{Q}^{*,b})$ is replaced by $\mathbb{Q}^*(Y^{(o),b} \neq Y^{(o)})$.

In Lemma 2 we showed that there exists $b_0 \geq 1$ such that if $b' > b \geq b_0$ then

$$\limsup_{n\to\infty} |\Lambda_n|^{-1} \log \mathbb{P}(\#\{X_i \in \Lambda_n \colon Y^{(i),b'} \neq Y^{(i),b}\} > \varepsilon|\Lambda_n|) \leq -K.$$

Hence, the LDP from Lemma 4 yields

$$\inf_{\mathbb{Q}_0\in\mathcal{P}_\theta,\, \mathbb{Q}_0^{*,b,b'}(Y^{(o),b'}\neq Y^{(o),b})>\varepsilon} h(\mathbb{Q} \mid \mathbb{P}) > K.$$

In particular, if $h(\mathbb{Q} \mid \mathbb{P}) \leq K$ then $\mathbb{Q}^{*,b',b}(Y^{(o),b'} \neq Y^{(o),b}) \leq \varepsilon$, as required. □

3. Proof of Theorem 2

The difficulty in proving Theorem 2 is that the connectable receivers associated with the origin are not stationary, so that we cannot use LDPs for the stationary empirical field directly. Therefore, we first consider a more general stationary marked point process from which the connectable receivers can be reproduced by an application of the contraction principle. Since we need a path-space LDP for this stationary marked point process, we proceed as in the classical proof of Mogulskii's theorem [10, Theorem 5.3.1] and use the Dawson–Gärtner technique to deduce the path-space LDP from the finite-dimensional marginals.

In order to define a suitable auxiliary stationary marked point process, we consider the random measure

$$M^t(\cdot) = \lambda_{\mathrm{R}} \int_{\Lambda'_t} \int_0^\infty \mathbf{1}\{\cdot\}\nu_y(\mathrm{d}s)\,\mathrm{d}y,$$

where $\nu_y([0, s]) = \Gamma(s, y) = \mathbb{E}(q(sI(y)^{-1}) \mid X)$; see also (1). Then, we let $Z^t = \{(Y_j, S_j)\}$ denote a Cox process with this random intensity measure. Writing

$$\Lambda_t(\xi^1, \ldots, \xi^d) = t^{-\beta} \prod_{i=1}^{d} \biggl[-\frac{|\Lambda'_1|^{1/d}}{2}, \xi^i\biggr],$$

define the two-parameter field $Y^{*,t} = \{Y^{*,t}(x, s)\}_{(x,s)\in\Lambda'_1\times[0,\infty)}$ by $Y^{*,t}(x, s) = Z^t(\Lambda_t(x) \times (0, s])$. In particular, for any fixed $(x, s) \in \Lambda'_1 \times [0, \infty)$, conditioned on X the random variable $Y^{*,t}(x, s)$ is Poisson distributed with parameter $\lambda_{\mathrm{R}} \int_{\Lambda_t(x)} \Gamma(s, y)\,\mathrm{d}y$. Moreover, $Y^{*,t}$ is a random variable with values in L_{inc}, the space of $[0, \infty)$-valued, bounded, and coordinate-wise increasing functions on $\Lambda'_1 \times [0, \infty)$.

We set $\mu_{\mathbb{Q}}(s) = \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(s, o))$ and note that the derivative $(\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s)$ exists since q is differentiable and $(\mathrm{d}/\mathrm{d}s)q(s)$ is Lipschitz continuous with Lipschitz constant N.

Similar to [10, Section 5.3], we introduce the notion of absolute continuity for increasing functions $F \colon \Lambda'_1 \times [0, \infty) \to [0, \infty)$, $(x, s) \mapsto F(x, s)$. For the convenience of the reader, we reproduce some of these definitions and observations. We define F as an additive set-function on the set of cubes. More precisely, for $\Lambda = (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_d, b_d] \times (a_{d+1}, b_{d+1}]$, we will sometimes write $F(\Lambda) := \sum_u \sigma(u)F(u)$ with $\sigma(u) \colon = (-1)^\rho$, where $\rho = \#\{k \colon u_k = a_k\}$ and the summation extends over all corners u of Λ; see [18, Chapter 3].

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

It follows from Carathéodory's extension theorem, that any right-continuous $F \in L_{\mathrm{inc}}$ induces a unique measure μ_F on $\Lambda_1' \times [0, \infty)$ with the Borel sigma-algebra satisfying

$$\mu_F\biggl(\prod_{i=1}^{d}\biggl[-\frac{|\Lambda_1'|^{1/d}}{2}, \xi^i\biggr] \times [0, s]\biggr) = F(\xi^1, \dots, \xi^d, s)$$

for any $s \geq 0$ and $(\xi^1, \dots, \xi^d) \in \Lambda_1'$; see [18, Theorem 3.25].

The function F is called absolutely continuous if F is right-continuous and μ_F is absolutely continuous with respect to the Lebesgue measure on $\Lambda_1' \times [0, \infty)$. We write $\partial F / \partial x \partial s$ for its Radon–Nikodym derivative. Let

$$\mathrm{AC}_0^1 := \biggl\{F \colon F \text{ is absolutely continuous, } F(x, 0) = 0 \text{ and } F\biggl(-\frac{|\Lambda_1'|^{1/d}}{2}, \xi^2, \dots, \xi^d, s\biggr) = F\biggl(\xi^1, -\frac{|\Lambda_1'|^{1/d}}{2}, \dots, \xi^d, s\biggr) = \cdots = F\biggl(\xi^1, \dots, \xi^{d-1}, -\frac{|\Lambda_1'|^{1/d}}{2}, s\biggr) = 0\biggr\}.$$

Later, we derive Theorem 2 by the contraction principle from the following result.

Proposition 2. *The random fields $\{|\Lambda_t'|^{-1} Y^{*,t}(\cdot, \cdot)\}_{t<1}$ satisfy an LDP in the topology of pointwise convergence with rate $|\Lambda_t'|$ and good rate function given by*

$$\mathcal{I}(F) = \begin{cases} \displaystyle\int_{\Lambda_1'} \mathcal{I}^*\biggl(\frac{\partial F}{\partial y \partial s}(y, \cdot)\biggr)\,\mathrm{d}y & \text{if } F \in AC_0^1, \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathcal{I}^*(g) = \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \int_0^\infty h(g(s) \mid (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s))\,\mathrm{d}s$.

3.1. Finite-dimensional result

In order to apply the Dawson–Gärtner theorem [10, Theorem 4.6.1], we first derive the finite-dimensional LDPs.

Proposition 3. *Let $-|\Lambda_1'|^{1/d}/2 = \xi_0 < \xi_1 < \cdots < \xi_k \leq |\Lambda_1'|^{1/d}/2$ and $0 = s_0 < s_1 < \cdots < s_r$. Furthermore, set $\Xi = \{\xi_0, \xi_1, \dots, \xi_k\}$ and $S = \{s_0, s_1, \dots, s_r\}$. Then, the random variables $\{(|\Lambda_t'|^{-1} Y^{*,t}(x, s))_{(x,s) \in \Xi^d \times S}\}_{t<1}$ satisfy an LDP with rate $|\Lambda_t'|$. Writing $\Delta\mu_{\mathbb{Q}}(s_i) = \mu_{\mathbb{Q}}(s_i) - \mu_{\mathbb{Q}}(s_{i-1})$, the good rate function is given by*

$$\mathcal{I}_{\Xi,S}(F) = \sum_{x \in \Xi^d} |\Lambda_\Xi(x)| \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h\biggl(\frac{1}{|\Lambda_\Xi(x)|} F(\Lambda_\Xi(x) \times (s_{i-1}, s_i]) \mid \Delta\mu_{\mathbb{Q}}(s_i)\biggr),$$

where $\Lambda_\Xi(\xi_{i_1}, \dots, \xi_{i_d}) = \prod_{j=1}^{d} (\xi_{i_j}, \xi_{i_j+1}]$ with $(\xi_{i_1}, \dots, \xi_{i_d}) \in \Xi^d$.

The basic idea of the proof for Proposition 3 is to apply the LDP for the stationary empirical field [16, Theorem 3.1]. However, in order to cast our problem into a suitable framework, we first have to perform a truncation and a periodization step.

3.1.1. *Truncation of the path-loss function.* In a first step, we show that truncation of the path-loss function gives an exponentially good approximation. Let $b \geq 1$, $s' \geq s \geq 0$, and $x, x' \in \Lambda_1$ be such that all coordinates of $x' - x$ are positive. Then, we let $Y^{*,b,t}(x, x', s, s')$ denote a random variable that conditioned on the independently marked Poisson particle process X

is Poisson distributed with parameter $\lambda_{\mathrm{R}}\int_{\Lambda_t(x,x')}\Gamma^b(s',y)-\Gamma^b(s,y)\,\mathrm{d}y$, where $\Gamma^b(s,y)=\mathbb{E}(sI^b(y)^{-1}\mid X)$,

$$\Lambda_t(x,x')=t^{-\beta}\Lambda(x,x')=t^{-\beta}\prod_{i=1}^{d}(\pi_k(x),\pi_k(x')],$$

and $\pi_k\colon\mathbb{R}^d\to\mathbb{R}$ denotes the projection onto the kth coordinate. From now on let us again assume that the support of the power variables is contained in $[0,s_{\min}^{-1}]$.

Lemma 5. *Let $b\geq 1$, $s'\geq s>0$, and $x,x'\in\Lambda'_1$. Then, $\{Y^{*,b,t}(x,x',s,s')\}_{b\geq1,t<1}$ are exponentially good approximations of $\{Y^{*,t}(\Lambda_t(x,x')\times(s,s'])\}_{t<1}$.*

Proof. Conditioned on X, the random variable $|Y^{*,b,t}(x,x',s,s')-Y^{*,t}(x,x',s,s')|$ is stochastically dominated by a Poisson distributed random variable with parameter

$$H_b=\lambda_{\mathrm{R}}\int_{\Lambda_t(x,x')}\Gamma^b(s',y)-\Gamma(s',y)+\Gamma^b(s,y)-\Gamma(s,y)\,\mathrm{d}y.$$

Hence, using the Laplace transform of Poisson random variables, for any $a\geq 1$, the exponential moments of $a|Y^{*,b,t}(x,x',s,s')-Y^{*,t}(x,x',s,s')|$ are bounded above by

$$\mathbb{E}\exp(a|Y^{*,b,t}(x,x',s,s')-Y^{*,t}(x,x',s,s')|)\leq\mathbb{E}\exp((\mathrm{e}^a-1)H_b).$$

Now, similar to the proof of Lemma 2, H_b can be bounded above by

$$\lambda_{\mathrm{R}}N(s+s')s_{\min}^{-2}w^{-2}\int_{\Lambda_t(x)}\sum_{i\geq1}\ell(|X_i-y|)-\ell_b(|X_i-y|)\,\mathrm{d}y.$$

Using this observation, we can now conclude as in Lemma 2 since

$$\begin{aligned}&\mathbb{E}\exp(a|Y^{*,b,t}(x,x',s,s')-Y^{*,t}(\Lambda_t(x,x')\times(s,s'])|)\\&\quad\leq\mathbb{E}\exp\biggl((\mathrm{e}^a-1)\lambda_{\mathrm{R}}N(s+s')s_{\min}^{-2}w^{-2}\int_{\Lambda_t(x,x')}\sum_{i\geq1}\ell(|X_i-y|)-\ell_b(|X_i-y|)\,\mathrm{d}y\biggr).\end{aligned}$$

This completes the proof. □

3.1.2. *Periodization of the integration domain.* Next, we show that replacing the quantity $Y^{*,b,t}(x,x',s,s')$ by a periodized variant is exponentially equivalent. To be more precise, let $b\geq1$, $s'\geq s\geq0$, and $x,x'\in\Lambda'_1$ be such that all coordinates of $x'-x$ are positive. Then, $X^{\mathrm{per},t}$ denotes the periodization of $X\cap\Lambda_t(x,x')$, i.e.

$$X^{\mathrm{per},t}=\bigcup_{z\in\mathbb{Z}^d}(|\Lambda_t(x,x')|^{1/d}z+X\cap\Lambda_t(x,x')).$$

As in Lemma 5, we let $Y^{*,\mathrm{per},b,t}(x,x',s,s')$ denote a random variable that conditioned on X is Poisson distributed with parameter $\lambda_{\mathrm{R}}\int_{\Lambda_t(x,x')}\Gamma^{\mathrm{per},b}(s',y)-\Gamma^{\mathrm{per},b}(s,y)\,\mathrm{d}y$. Here, $\Gamma^{\mathrm{per},b}(s,y)=\mathbb{E}(q(sI^{\mathrm{per},b}(y)^{-1})\mid X^{\mathrm{per},t})$ and $I^{\mathrm{per},b}(y)$ is the interference at y in the periodized configuration computed using truncated path-loss functions.

Lemma 6. *The random variables $\{Y^{*,\mathrm{per},b,t}(x,x',s,s')\}_{t<1}$ are exponentially equivalent to the random variables $\{Y^{*,b,t}(x,x',s,s')\}_{t<1}$.*

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

Proof. Since we consider truncated interferences, we have $I^{\mathrm{per},b}(y) = I^b(y)$ for all $y \in \Lambda_t^-(x, x')$, where $\Lambda_t^-(x, x')$ denotes the subset of all $y \in \Lambda_t(x, x')$ such that $B_b(y) \subset \Lambda_t(x, x')$. In particular, $|Y^{*,\mathrm{per},b,t}(x, x', s, s') - Y^{*,b,t}(x, x', s, s')|$ is stochastically dominated by a Poisson random variable with parameter $2\lambda_{\mathrm{R}}|\Lambda_t(x, x') \setminus \Lambda_t^-(x, x')|$. Now, using the Poisson concentration property we can conclude as in Lemma 3. □

3.1.3. *Application of an LDP for stationary empirical fields.* We have seen that truncating the interference and considering a periodization does not have an effect on $\{Y^{*,t}(x, x', s, s')\}_{t<1}$ in the LDP asymptotics. Now, we derive an LDP after these modifications have been implemented. We set $\mu_{\mathbb{Q}}^b(s) = \mathbb{Q}(\Gamma^b(s, o))$.

Proposition 4. *The random variables* $\{|\Lambda_t'|^{-1}\#Y^{*,\mathrm{per},b,t}(x, x', s, s')\}_{\tau<1}$ *satisfy an LDP with rate* $|\Lambda_t'|$ *and good rate function*

$$\mathcal{I}_{b,N}^{x,x',s,s'}(a) = |\Lambda(x, x')| \inf_{\mathbb{Q}\in\mathcal{P}_\theta} \biggl(h(\mathbb{Q} \mid \mathbb{P}) + h\biggl(\frac{a}{|\Lambda(x, x')|} \biggm| \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s)\biggr)\biggr).$$

Let us recall from [10, Equations 1.2.12 and 1.2.13] that if the random variable considered in an LDP is measurable with respect to the Borel σ-algebra on the underlaying topological space, then the proof of the upper and lower bound can be carried out directly for closed and open sets, respectively. We use this in the sequel without further mention.

We prepare the proof of Proposition 4 by a lemma. First, we note that [16, Theorem 3.1] gives the following auxiliary result, where we set

$$M_{\mathrm{av},t} = M_{\mathrm{av},t}(x, x', s, s') = \lambda_{\mathrm{R}}|\Lambda_t(x, x')|^{-1} \int_{\Lambda_t(x,x')} \Gamma^{\mathrm{per},b}(s', y) - \Gamma^{\mathrm{per},b}(s, y)\,\mathrm{d}y.$$

Lemma 7. *Let F and G be compact and open subsets of* $[0, \infty)$*, respectively. Then,*

$$\begin{aligned}\limsup_{t\to 0} \frac{1}{|\Lambda_t(x, x')|} &\log \mathbb{E} \exp\Bigl(-|\Lambda_t(x, x')| \inf_{a\in F} h(a \mid M_{\mathrm{av},\tau})\Bigr) \\ &\le -\inf_{\mathbb{Q}\in\mathcal{P}_\theta,\, a\in F} h(\mathbb{Q} \mid \mathbb{P}) + h(a \mid \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s)),\end{aligned}$$

and

$$\begin{aligned}\liminf_{t\to 0} \frac{1}{|\Lambda_t(x, x')|} &\log \mathbb{E} \exp\Bigl(-|\Lambda_t(x, x')| \inf_{a\in G} h(a \mid M_{\mathrm{av},\tau})\Bigr) \\ &\ge -\inf_{\mathbb{Q}\in\mathcal{P}_\theta,\, a\in G} h(\mathbb{Q} \mid \mathbb{P}) + h(a \mid \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s)).\end{aligned}$$

Proof. In order to apply [16, Theorem 3.1], we need to check that the functions $\mathbb{Q} \mapsto \inf_{a\in F} h(a \mid \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s))$ and $\mathbb{Q} \mapsto \inf_{a\in G} h(a \mid \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s))$ are lower- and upper-semicontinuous, respectively. First, $\mathbb{Q} \mapsto \mu_{\mathbb{Q}}^b(s') - \mu_{\mathbb{Q}}^b(s)$ is continuous in the $\tau_{\mathcal{L}}$-topology, since $\Gamma^b(\cdot, o)$ only depends on X via $X \cap B_b(o)$. Now, we conclude by observing that $a' \mapsto \inf_{a\in F} h(a \mid a')$ is lower-semicontinuous as pointwise infimum of a two-parameter lower-semicontinuous function over a compact set and $a' \mapsto \inf_{a\in G} h(a \mid a')$ is upper-semicontinuous as infimum over a family of continuous functions. □

Now, we can proceed with the proof of Proposition 4.

Proof of Proposition 4. The upper bound for compact F is an immediate consequence of Lemma 7, since [21, Lemma 1.2] implies that

$$\mathbb{P}(|\Lambda_t'|^{-1}Y^{*,\mathrm{per},b,t}(x,x',s,s')\in F)\le \mathbb{E}\exp\bigg(-|\Lambda_t(x,x')|\inf_{a\in F}h\bigg(\frac{a}{|\Lambda(x,x')|}\;\bigg|\;M_{\mathrm{av},\tau}\bigg)\bigg).$$

The proof of the lower bound is more involved. First, we may assume that G is an interval, i.e. $G=[0,\gamma)$ or $G=(\gamma_-,\gamma_+)$ for some $\gamma,\gamma_-,\gamma_+>0$. Next, introduce the function $f(k)$ by $f(0)=1$ and $f(k)=\mathrm{e}^{-1/12k}(\sqrt{2\pi k})^{-1}$ for $k\ge 1$, and set $G^t=\mathbb{Z}\cap(|\Lambda_t'|G)$. Then, by [21, Lemma 1.3], $\mathbb{P}(|\Lambda_t'|^{-1}Y^{*,\mathrm{per},b,t}(x,x',s,s')\in G)$ is bounded below by

$$\begin{aligned}&\mathbb{E}\exp\Big(-\inf_{k\in G^t}-\log f(k)+h(k|\Lambda_t(x,x')|M_{\mathrm{av},t})\Big)\\&\quad\ge\mathbb{E}\exp\Big(-|\Lambda_t'|(|\Lambda_t'|^{-1/2}+|\Lambda(x,x')|\inf_{k\in G^t}M_{\mathrm{av},t}h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1}))\Big),\end{aligned}$$

where $h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1})=h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1}\mid 1)$ is abbreviated notation. Now, we distinguish between the cases where G contains 0 and where it does not. We claim that if $G=[0,\gamma)$ and $\varepsilon>0$, then $\inf_{k\in G^t}M_{\mathrm{av},t}h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1})$ is at most

$$\varepsilon+\inf_{g\in G}M_{\mathrm{av},t}h\bigg(\frac{g}{|\Lambda(x,x')|}M_{\mathrm{av},t}^{-1}\bigg),$$

provided that $t>0$ is sufficiently small. Once this claim is proven, Lemma 7 completes the proof of the lower bound for the $G=[0,\gamma)$ case. Let $\varepsilon>0$ be arbitrary. Then, under the event $M_{\mathrm{av},t}\le\varepsilon$, we deduce that $\inf_{k\in G^t}M_{\mathrm{av},t}h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\le M_{\mathrm{av},t}\le\varepsilon$. On the other hand, if $M_{\mathrm{av},t}\ge\varepsilon$ then, for every $g\in G$,

$$|g|\Lambda(x,x')|^{-1}M_{\mathrm{av},t}^{-1}-k(g)|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1}|\le|\Lambda_t(x,x')|^{-1}\varepsilon^{-1},$$

where $k(g)\ge 1$ is chosen as the element of $\mathbb{Z}\cap(|\Lambda_t'|G)$ such that $k(g)|\Lambda_t(x,x')|^{-1}$ minimizes the distance to $g|\Lambda(x,x')|^{-1}$. In particular, uniform continuity of $h(\cdot)$ on the interval $[0,\gamma|\Lambda(x,x')|^{-1}\varepsilon^{-1}]$ implies that

$$\inf_{k\in G^t}M_{\mathrm{av},t}h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\le\varepsilon+\inf_{g\in G}M_{\mathrm{av},t}h(g|\Lambda(x,x')|^{-1}M_{\mathrm{av},t}^{-1})$$

for all sufficiently small $t>0$. Finally, we deal with the case when $G=(\gamma_-,\gamma_+)$ and observe that if $M_{\mathrm{av},t}\ge\varepsilon$ then we can conclude as before. To be more precise,

$$\begin{aligned}&\mathbb{E}\exp\Big(-|\Lambda_t(x,x')|\inf_{k\in G^t}M_{\mathrm{av},t}h(k|\Lambda_t(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\Big)\\&\quad\ge\mathbb{E}\exp\Big(-|\Lambda_t(x,x')|\Big(-\varepsilon+\inf_{g\in G}M_{\mathrm{av},t}h(g|\Lambda(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\Big)\Big)\\&\qquad-\mathbb{E}\mathbf{1}\{M_{\mathrm{av},t}\le\varepsilon\}\exp\Big(-|\Lambda_t(x,x')|\Big(-\varepsilon+\inf_{g\in G}M_{\mathrm{av},t}h(g|\Lambda(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\Big)\Big).\end{aligned}$$

Now, for any $K\ge 1$, there exists $\varepsilon>0$ such that $M_{\mathrm{av},t}h(\gamma_-|\Lambda(x,x')|^{-1}M_{\mathrm{av},t}^{-1})\ge K$ if $M_{\mathrm{av},t}\le\varepsilon$. In particular, we may complete the proof of the lower bound by noting that

$$\exp\Big(-|\Lambda_t(x,x')|\Big(-\varepsilon+\inf_{g\in G}M_{\mathrm{av},t}h(gM_{\mathrm{av},t}^{-1})\Big)\Big)\le\exp(-|\Lambda_t(x,x')|(-\varepsilon+K)).$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

Since $Y^{*,\mathrm{per},b,t}(x, x', s, s')$ is stochastically dominated by a Poisson random variable with parameter $\lambda_{\mathrm{R}}|\Lambda_t(x, x')|$, the random variables $\{|\Lambda'_t|^{-1}Y^{*,\mathrm{per},b,t}(x, x', s, s')\}_{t<1}$ are exponentially tight. This implies both goodness of the rate function and the full LDP. □

Next, using Lemma 4, we derive an LDP for the finite-dimensional marginals of $Y^{*,\mathrm{per},b,t}(\cdot,\cdot)$. In order to state this precisely, it is convenient to introduce some notation. Let $-|\Lambda'_1|^{1/d}/2 = \xi_0 < \xi_1 < \cdots < \xi_k \le |\Lambda'_1|^{1/d}/2$ and $0 = s_0 < s_1 < \cdots < s_r$. Then, for $x = (\xi_{i_1}, \ldots, \xi_{i_d})$ and $s = s_i$, we set $x_{+,\Xi} = (\xi_{i_1+1}, \ldots, \xi_{i_d+1})$ and $s_{+,S} = s_{i+1}$, where we use the conventions $\xi_{k+1} = |\Lambda'_1|^{1/d}/2$ and $s_{\ell+1} = \infty$.

Corollary 2. *Let* $-|\Lambda'_1|^{1/d}/2 = \xi_0 < \xi_1 < \cdots < \xi_k \le |\Lambda'_1|^{1/d}/2$ *and* $0 = s_0 < s_1 < \cdots < s_\ell$. *Furthermore, set* $\Xi = \{\xi_0, \xi_1, \ldots, \xi_k\}$ *and* $S = \{s_0, s_1, \ldots, s_r\}$. *Then, the random vectors* $\{(|\Lambda'_t|^{-1}Y^{*,b,t}(x, x_{+,\Xi}, s, s_{+,S}))_{(x,s)\in\Xi^d\times S}\}_{t<1}$ *satisfy an LDP with rate* $|\Lambda'_t|$ *and good rate function*

$$\mathcal{I}^b_{\Xi,S}((a_z)_{z\in\Xi^d\times S}) = \sum_{x\in\Xi^d} |\Lambda(x, x_{+,\Xi})| \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h\biggl(\frac{a_{x,s_i}}{|\Lambda(x, x_{+,\Xi})|} \,\bigg|\, \Delta\mu^b_{\mathbb{Q}}(s_i)\biggr).$$

Proof. First, we observe that $(Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi}))_{x\in\Xi^d}$ defines a family of independent random vectors, where we set $Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi}) = (Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi}, s, s_{+,S}))_{s\in S}$. Indeed, this is a consequence of the independence property of the Poisson point process X since by the definition of the periodization, $Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi})$ depends on X only via $X \cap \Lambda_t(x, x_{+,\Xi})$. Hence, if, for any fixed $x \in \Xi^d$, we can establish an LDP for $Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi})$ with a certain good rate function, then [10, Exercise 4.2.7] allows us to deduce that the collection $(Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi}))_{x\in\Xi^d}$ satisfies an LDP with good rate function given by the sum of the individual ones. If $|S| = 1$ then the LDP for $Y^{*,\mathrm{per},b,t}(x, x_{+,\Xi})$ is precisely the result of Lemma 4, and an inspection of its proof shows that it also extends to the case of general finite S.

In Lemma 5 we have seen that periodization replaces $\{Y^{*,b,t}(x, x', s, s')\}_{\tau<1}$ by exponentially equivalent random variables. Hence, by applying [10, Theorem 4.2.13] we are able to complete the proof. □

In order to deduce Proposition 3 from Corollary 2, we need to undo the truncation approximations. Before we start with the proof of Proposition 3, it is convenient to derive certain continuity properties of $\mu^b_{\mathbb{Q}}$ and $(\mathrm{d}/\mathrm{d}s)\mu^b_{\mathbb{Q}}(s)$ with respect to b and $\mathbb{Q}$. The technique of proof is similar to the one used in the proof of Theorem 1.

Lemma 8. *Let* $\varepsilon, K > 0$, *and* $s \ge 0$ *be arbitrary. Then, there exists* $b \ge 1$ *such that if* $\mathbb{Q} \in \mathcal{P}_\theta$ *satisfies* $h(\mathbb{Q} \mid \mathbb{P}) \le K$ *then* $\mu^b_{\mathbb{Q}}(s) - \mu_{\mathbb{Q}}(s) \le \varepsilon$ *and* $|(\mathrm{d}/\mathrm{d}s)\mu^b_{\mathbb{Q}}(s) - (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s)| \le \varepsilon$.

Proof. We first deal with the part of the statement not involving derivatives. Let $b' \ge b \ge 1$ be arbitrary. Proceeding as in Lemma 7, we see that $\{|\Lambda'_t|^{-1}\int_{\Lambda'_t}\Gamma^b(s, y) - \Gamma^{b'}(s, y)\,\mathrm{d}y\}_{t<1}$ satisfies an LDP with rate $|\Lambda'_t|$ and good rate function

$$a \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta,\ \mu^b_{\mathbb{Q}}(s)-\mu^{b'}_{\mathbb{Q}}(s)=a} h(\mathbb{Q} \mid \mathbb{P}).$$

In particular, the proof is completed once we show the existence of $b_0 \ge 1$ such that

$$\limsup_{t\to 0} |\Lambda'_t|^{-1} \log \mathbb{P}\biggl(\int_{\Lambda'_t} |\Lambda'_t|^{-1}\Gamma^b(s, y) - \Gamma^{b'}(s, y)\,\mathrm{d}y > \varepsilon\biggr) \le -K$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

for all $b' \geq b \geq b_0$. Note that

$$\Gamma^b(s, y) - \Gamma^{b'}(s, y) \leq Nss_{\min}^{-2} w^{-2} \sum_{i \geq 1} \ell_{b'}(|y - X_i|) - \ell_b(|y - X_i|). \tag{3}$$

Using this result, we may now conclude as in Lemma 2, since the formula for the Laplace functional of a Poisson point process shows that, for any $a > 0$,

$$\begin{aligned} &\mathbb{E} \exp\biggl(a \int_{\Lambda_t'} \Gamma^b(s, y) - \Gamma^{b'}(s, y) \, \mathrm{d}y > \varepsilon\biggr) \\ &\quad \leq \exp\biggl(\lambda_{\mathrm{T}} \int_{\mathbb{R}^d} \exp\biggl(aNs_{\min}^{-2} w^{-2} \int_{\Lambda_t'} \ell(|x - y|) - \ell_b(|x - y|) \, \mathrm{d}y\biggr) - 1 \, \mathrm{d}x\biggr). \end{aligned}$$

Considering the part involving the derivatives, note that the derivative of $\mu_{\mathbb{Q}}(s)$ is given by $\lambda_{\mathrm{R}} \mathbb{Q}(I(o)^{-1} (\mathrm{d}/\mathrm{d}s) \Gamma(s, o))$. Essentially, this means replacing in the above arguments the expression $q(sI(y)^{-1})$ by $I(y)^{-1} (\mathrm{d}/\mathrm{d}s) q(sI(y)^{-1})$. We conclude by observing that this specific form only comes into play in the estimate (3), which can be replaced by

$$\begin{aligned} &\biggl| I^b(y)^{-1} \frac{\mathrm{d}}{\mathrm{d}s} q(sI^b(y)^{-1}) - I^{b'}(y)^{-1} \frac{\mathrm{d}}{\mathrm{d}s} q(sI^{b'}(y)^{-1}) \biggr| \\ &\quad \leq I^b(y)^{-1} \biggl| \frac{\mathrm{d}}{\mathrm{d}s} q(sI^b(y)^{-1}) - \frac{\mathrm{d}}{\mathrm{d}s} q(sI^{b'}(y)^{-1}) \biggr| \\ &\qquad + \frac{\mathrm{d}}{\mathrm{d}s} q(sI^{b'}(y)^{-1}) |I^b(y)^{-1} - I^{b'}(y)^{-1}| \\ &\quad \leq (Nss_{\min}^{-2} w^{-3} + Nw^{-2} s_{\min}^{-2}) \sum_{i \geq 1} (\ell_{b'}(|y - X_i|) - \ell_b(|y - X_i|)). \end{aligned}$$ □

Corollary 3. *Let $K > 0$ and $s \geq 1$ be arbitrary. Then, in the $\tau_{\mathcal{L}}$-topology,*

- *as $b \to \infty$, the functions $\mathbb{Q} \mapsto \mu_{\mathbb{Q}}^b(s)$ converge to $\mu_{\mathbb{Q}}(s)$ uniformly in $\{\mathbb{Q} : h(\mathbb{Q} \mid \mathbb{P}) \leq K\}$. In particular, $\mathbb{Q} \mapsto \mu_{\mathbb{Q}}(s)$ is continuous on $\{\mathbb{Q} : h(\mathbb{Q} \mid \mathbb{P}) \leq K\}$;*
- *as $b \to \infty$, the functions $\mathbb{Q} \mapsto (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}^b(s)$ converge to $(\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s)$ uniformly in $\{\mathbb{Q} : h(\mathbb{Q} \mid \mathbb{P}) \leq K\}$. In particular, $\mathbb{Q} \mapsto (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s)$ is continuous on $\{\mathbb{Q} : h(\mathbb{Q} \mid \mathbb{P}) \leq K\}$;*
- *if $b_n \to \infty$, $\mathbb{Q}_n \to \mathbb{Q}$, and $\limsup_{n \to \infty} h(\mathbb{Q}_n \mid \mathbb{P}) \leq K$, then $\mu_{\mathbb{Q}_n}^{b_n}(s) \to \mu_{\mathbb{Q}}(s)$.*

Proof. Since the first two items are an immediate consequence of Lemma 8, we only deal with the last item. Here, we may conclude, as before, by using the decomposition

$$|\mu_{\mathbb{Q}}(s) - \mu_{\mathbb{Q}_n}^{b_n}(s)| \leq |\mu_{\mathbb{Q}}(s) - \mu_{\mathbb{Q}}^b(s)| + |\mu_{\mathbb{Q}}^b(s) - \mu_{\mathbb{Q}_n}^b(s)| + |\mu_{\mathbb{Q}_n}^{b_n}(s) - \mu_{\mathbb{Q}_n}^b(s)|.$$ □

Now, we have completed all preparations for the proof of Proposition 3.

Proof of Proposition 3. In Lemma 5 we have seen that truncation leads to an exponentially good approximation. Therefore, combining Corollary 2 with [10, Theorem 4.2.16] shows that the random vector $\{(|\Lambda_t'|^{-1} Y^{*,t}(x, s))_{(x,s) \in \Xi^d \times S}\}_{t<1}$ satisfies a weak LDP with rate function

$$\mathcal{I}_{\Xi,S}'((a_{x,s})_{(x,s)}) = \sup_{m \geq 1} \liminf_{b \to \infty} \inf_{\{(a_{x,s}')_{(x,s)} : \, |(a_{x,s}')_{(x,s)} - (a_{x,s})_{(x,s)}|_\infty \leq m^{-1}\}} \mathcal{I}_{\Xi,S}^b((a_{x,s}')_{(x,s)}).$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

Since the random vectors $\{(|\Lambda_t'|^{-1}Y^{*,t}(x,s))_{(x,s)\in\Xi^d\times S}\}_{\tau<1}$ are exponentially tight, the proof is completed once we show that $\mathcal{I}'_{\Xi,S}((a_{x,s})_{(x,s)}) = \mathcal{I}_{\Xi,S}((a_{x,s})_{(x,s)})$, where

$$\mathcal{I}_{\Xi,S}((a_{x,s})_{(x,s)}) = \sum_{x\in\Xi^d} |\Lambda(x,x_{+,\Xi})| \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q}\mid\mathbb{P}) + \sum_{i=1}^r h\bigg(\frac{a_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu_{\mathbb{Q}}(s_i)\bigg).$$

First, we note that $q(sI^b(o)^{-1})$ is decreasing in b and converges to $q(sI(o)^{-1})$ as $b\to\infty$. Hence, for any $\mathbb{Q}\in\mathcal{P}_\theta$ and $x\in\Xi^d$,

$$\lim_{b\to\infty}\sum_{i=1}^r h\bigg(\frac{a_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^b_{\mathbb{Q}}(s_i)\bigg) = \sum_{i=1}^r h\bigg(\frac{a_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu_{\mathbb{Q}}(s_i)\bigg).$$

In particular, $\mathcal{I}'_{\Xi,S}((a_{x,s})_{(x,s)}) \le \mathcal{I}_{\Xi,S}((a_{x,s})_{(x,s)})$. Conversely, fix $\delta>0$ and $x\in\Xi^d$. Then, for $m\ge1$, choose a sequence, $(b_{m,n})_{n\ge1}$, such that $\lim_{n\to\infty} b_{m,n}=\infty$ and write

$$\begin{aligned}
&\lim_{n\to\infty}\inf_{\{\mathbb{Q}\in\mathcal{P}_\theta,\,(a'_{x,s})_{(x,s)}:\,|(a'_{x,s})_s-(a_{x,s})_s|_\infty\le m^{-1}\}} h(\mathbb{Q}\mid\mathbb{P}) + \sum_{i=1}^r h\bigg(\frac{a'_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b_{m,n}}_{\mathbb{Q}}(s_i)\bigg)\\
&\quad= \liminf_{b\to\infty}\inf_{\{\mathbb{Q}\in\mathcal{P}_\theta,\,(a'_{x,s})_{(x,s)}:\,|(a'_{x,s})_s-(a_{x,s})_s|_\infty\le m^{-1}\}} h(\mathbb{Q}\mid\mathbb{P})\\
&\qquad+\sum_{i=1}^r h\bigg(\frac{a'_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b}_{\mathbb{Q}}(s_i)\bigg).
\end{aligned}$$

Next, for each $m,n\ge1$, choose $\mathbb{Q}_{x,m,n}\in\mathcal{P}_\theta$, and for each $m,n\ge1$, and $s\in S$ choose $a'_{x,s,m,n}\in[a_{x,s}-1/m,a_{x,s}+1/m]$ such that

$$\begin{aligned}
&\inf_{\{\mathbb{Q}^*\in\mathcal{P}_\theta,\,(a^*_{x,s})_s:\,|(a^*_{x,s})_s-(a_{x,s})_s|_\infty\le m^{-1}\}} h(\mathbb{Q}^*\mid\mathbb{P}) + \sum_{i=1}^r h\bigg(\frac{a^*_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b_{m,n}}_{\mathbb{Q}^*}(s_i)\bigg)\\
&\quad\ge -\delta + h(\mathbb{Q}_{x,m,n}\mid\mathbb{P}) + \sum_{i=1}^r h\bigg(\frac{a'_{x,s_i,m,n}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b_{m,n}}_{\mathbb{Q}_{x,m,n}}(s_i)\bigg).
\end{aligned}$$

If $\limsup_{n\to\infty} h(\mathbb{Q}_{x,m,n}\mid\mathbb{P})=\infty$ then

$$\begin{aligned}
&\lim_{n\to\infty}\inf_{\{\mathbb{Q}\in\mathcal{P}_\theta,\,(a'_{x,s})_{(x,s)}:\,|(a'_{x,s})_s-(a_{x,s})_s|_\infty\le m^{-1}\}} h(\mathbb{Q}\mid\mathbb{P}) + \sum_{i=1}^r h\bigg(\frac{a'_{x,s_i}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b_{m,n}}_{\mathbb{Q}}(s_i)\bigg)\\
&\quad=\infty,
\end{aligned}$$

which is certainly at least as large as $\mathcal{I}_{\Xi,S}((a_{x,s})_{(x,s)})$. Otherwise, after passing to a subsequence, we may assume that $\lim_{n\to\infty}\mathbb{Q}_{x,m,n}=\mathbb{Q}_{x,m}$ for some $\mathbb{Q}_{x,m}\in\mathcal{P}_\theta$ by sequential compactness. Furthermore, we may also assume, for each $1\le i\le r$, that $\lim_{n\to\infty}a'_{x,s_i,m,n}=a'_{x,s_i,m}$ for some $a'_{x,s_i,m}\in[a_{x,s}-1/m,a_{x,s}+1/m]$. In particular, lower semicontinuity of h implies that $\liminf_{n\to\infty}h(\mathbb{Q}_{x,m,n}\mid\mathbb{P})\ge h(\mathbb{Q}_{x,m}\mid\mathbb{P})$. Moreover, by Corollary 3, $\lim_{n\to\infty}\Delta\mu^{b_{m,n}}_{\mathbb{Q}_{x,m,n}}(s_i)=\Delta\mu_{\mathbb{Q}_{x,m}}(s_i)$. Hence, by another application of lower semicontinuity, we have

$$\liminf_{n\to\infty} h\bigg(\frac{a'_{x,s_i,m,n}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu^{b_{m,n}}_{\mathbb{Q}_{x,m,n}}(s_i)\bigg) \ge h\bigg(\frac{a'_{x,s_i,m}}{|\Lambda(x,x_{+,\Xi})|}\,\bigg|\,\Delta\mu_{\mathbb{Q}_{x,m}}(s_i)\bigg).$$

Arguing as above, we may assume that $\mathbb{Q}_{x,m}$ converges to some $\mathbb{Q}_x$ as $m \to \infty$. In order to conclude the proof of the proposition, it therefore suffices to show that

$$\liminf_{m\to\infty} h\biggl(\frac{a'_{x,s_i,m}}{|\Lambda(x, x_{+,\Xi})|} \biggm| \Delta\mu_{\mathbb{Q}_{x,m}}(s_i)\biggr) \geq h\biggl(\frac{a_{x,s_i}}{|\Lambda(x, x_{+,\Xi})|} \biggm| \Delta\mu_{\mathbb{Q}_x}(s_i)\biggr).$$

A final application of lower semicontinuity completes the proof. □

3.2. Application of Dawson–Gärtner and identification of a rate function

In Proposition 3, we have shown that the finite-dimensional distributions of the random fields $\{Y^{*,t}\}_{t<1}$ satisfy an LDP. We have also identified the good rate function. Hence, the Dawson–Gärtner theorem [10, Theorem 4.6.1] implies that the random fields $\{Y^{*,t}\}_{t<1}$ satisfy an LDP with respect to the topology of pointwise convergence and that the good rate function is given by $\tilde{\mathcal{I}}(F) = \sup_{\Xi,S} \mathcal{I}_{\Xi,S}(F)$, where the supremum is over all finite $S \subset [0,\infty)$ and $\Xi \subset [-|\Lambda'_1|^{1/d}/2, |\Lambda'_1|^{1/d}/2]$. The proof of Proposition 2 now amounts to showing $\tilde{\mathcal{I}}(F) = \mathcal{I}(F)$. This is done using an adaptation of arguments appearing in the classical derivation of Mogulskii's theorem provided in [10, Theorem 5.3.1]. For the convenience of the reader, we provide some details.

Proof of Proposition 2. First assume that $F \in AC_0^1$ and let $f = \partial F/\partial x \partial s$ denote the density of F. By nonnegativity of $\mathcal{I}_{\Xi,S}(F)$, we can assume that $\xi_k = \Lambda'_1|^{1/d}/2$. Note that $\mathcal{I}_{\Xi,S}(F)$ is of the form $\sum_{x\in\Xi^d} |\Lambda_\Xi(x)| \inf_{\mathbb{Q}} f(F(x), \mathbb{Q})$, where f is convex in the pair $(F(x), \mathbb{Q})$ by linearity of $\mu_{\mathbb{Q}}$. Hence, also $G(F(x)) = \inf_{\mathbb{Q}} f(F(x), \mathbb{Q})$ is convex in $F(x)$ so that using Jensen's inequality, we have

$$\mathcal{I}_{\Xi,S}(F) \leq \int_{\Lambda'_1} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{k} h\biggl(\int_{s_{i-1}}^{s_i} f(x,s)\,\mathrm{d}s \mid \Delta\mu_{\mathbb{Q}}(s_i)\biggr)\,\mathrm{d}x.$$

Again by convexity of h and Jensen's inequality, we can further estimate

$$h\biggl(\int_{s_{i-1}}^{s_i} f(x,s)\,\mathrm{d}s \mid \Delta\mu_{\mathbb{Q}}(s_i)\biggr) \leq \int_{s_{i-1}}^{s_i} h\biggl(f(x,s) \biggm| \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}}(s)\biggr)\,\mathrm{d}s.$$

This proves $\tilde{\mathcal{I}}(F) \leq \mathcal{I}(F)$. For the other direction, we first consider the supremum over partitions Ξ and let some S-partition be fixed. The idea is to use a volume partition into equal subcubes with side length going to 0 as a lower bound. More precisely, let $\{\rho_k(l)\}_{l=1}^{k^d}$ denote the disjoint partition of Λ'_1 into cubes of volume $|\Lambda'_1|/k^d$ and equal side-length $\delta(k) = |\Lambda'_1|^{1/d}/k$. Then,

$$\begin{aligned}\mathcal{I}_{\Xi,S}(F) &\geq \liminf_{k\to\infty} \sum_{l=1}^{k^d} \frac{1}{k^d} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h(F(\rho_k(l) \times (s_{i-1}, s_i]) \mid \Delta\mu_{\mathbb{Q}}(s_i)) \\ &= \liminf_{k\to\infty} \int_{\Lambda'_1} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h(f_i^k(x) \mid \Delta\mu_{\mathbb{Q}}(s_i))\,\mathrm{d}x,\end{aligned}$$

where each $f_i^k(x)$ is constant on each of the cubes $\rho_k(l), l = 1, \dots, k^d$. By Lebesgue's theorem, $f_i^k(x) \to \int_{s_{i-1}}^{s_i} f(x,s)\,\mathrm{d}s$ for Lebesgue almost all x as k tends to ∞. Therefore, by Fatou's

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

lemma and the lower semicontinuity of the rate function,

$$\begin{aligned}\sup_{\Xi} \mathcal{I}_{\Xi,S}(F) &\geq \liminf_{k\to\infty} \int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h(f_i^k(x) \mid \Delta\mu_{\mathbb{Q}}(s_i))\,\mathrm{d}x \\ &\geq \int_{\Lambda_1'} \liminf_{k\to\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h(f_i^k(x) \mid \Delta\mu_{\mathbb{Q}}(s_i))\,\mathrm{d}x \\ &= \int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{r} h\biggl(\int_{s_{i-1}}^{s_i} f(x,s)\,\mathrm{d}s \mid \Delta\mu_{\mathbb{Q}}(s_i)\biggr)\,\mathrm{d}x \\ &= \mathcal{I}_S(F).\end{aligned}$$

For the supremum over S-partitions, we use the same approach and consider a partition of intervals $[0, k]$ for $k \in \mathbb{N}$ with constant mesh size $1/k$. Using Fatou's lemma, we have

$$\sup_S \mathcal{I}_S(F) \geq \int_{\Lambda_1'} \liminf_{k\to\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \sum_{i=1}^{k^2} \frac{1}{k} h\biggl(\frac{\int_{(i-1)/k}^{i/k} f(x,s)\,\mathrm{d}s}{1/k} \biggm| \frac{\Delta\mu_{\mathbb{Q}}(i/k)}{1/k}\biggr)\,\mathrm{d}x,$$

where we can look at $k\int_{(i-1)/k}^{i/k} f(x,s)\,\mathrm{d}s$ as a stepfunction f_x^k on $[0,k]$. Similarly, for $k[\mu_{\mathbb{Q}}(i/k) - \mu_{\mathbb{Q}}((i-1)/k)] = k\int_{(i-1)/k}^{i/k} (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s)\,\mathrm{d}s$ with $g_{\mathbb{Q}}^k$ on $[0,k]$, we have

$$\sum_{i=1}^{k^2} \frac{1}{k} h\biggl(k\int_{s_{i-1}}^{s_i} f(x,s)\,\mathrm{d}s \mid k\Delta\mu_{\mathbb{Q}}\biggl(\frac{i}{k}\biggr)\biggr) = \int_0^k h(f_x^k(r) \mid g_{\mathbb{Q}}^k(r))\,\mathrm{d}r.$$

Now fix $x \in \Lambda_1$ and $k \geq 1$ and let $\mathbb{Q}_k^x$ such that

$$\inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \int_0^k h(f_x^k(r) \mid f_{\mathbb{Q}}^k(r))\,\mathrm{d}r = h(\mathbb{Q}_k^x \mid \mathbb{P}) + \int_0^k h(f_x^k(r) \mid g_{\mathbb{Q}_k^x}^k(r))\,\mathrm{d}r,$$

which exists since lower-semicontinuous functions assume their minimum on compact sets. Let k_n be the subsequence such that the limit inferior becomes a limit and for simplicity write again k. Furthermore, we can assume that $\sup_k h(\mathbb{Q}_k^x \mid \mathbb{P}) < \infty$ for Lebesgue almost all x since otherwise there is nothing to show. Since $h(\cdot \mid \mathbb{P})$ has sequentially compact level sets there exists a cluster point $\mathbb{Q}_*^x$ of $(\mathbb{Q}_k^x)_{k\in\mathbb{N}}$ and by lower semicontinuity and Fatou's lemma, we have $\sup_S \mathcal{I}_S(F) \geq \int_{\Lambda_1'} h(\mathbb{Q}_*^x \mid \mathbb{P}) + \int_0^\infty \liminf_{k\to\infty} h(f_x^k(r) \mid g_{\mathbb{Q}_k^x}^k(r))\,\mathrm{d}r\,\mathrm{d}x$. Note that by Lebesgue's theorem for almost all $s \in [0,\infty)$, $\liminf_{k\to\infty} f_x^k(s) = f(x,s)$. Furthermore, $\liminf_{k\to\infty} g_{\mathbb{Q}_k^x}^k(s) = (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}_*^x}(s)$. Indeed, by the mean value theorem for $s \in ((i-1)/k, i/k)$ there exists $s' \in ((i-1)/k, i/k)$ such that $g_{\mathbb{Q}_k^x}^k(s) = (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}_k^x}(s')$ and

$$\biggl|g_{\mathbb{Q}_k^x}^k(s) - \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_*^x}(s)\biggr| \leq \biggl|\frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_k^x}(s') - \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_k^x}(s)\biggr| + \biggl|\frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_k^x}(s) - \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_*^x}(s)\biggr|.$$

The second summand on the right tends to 0 as k tends to ∞ by Corollary 3. For the first term, we have, by Lebesgue's theorem,

$$\biggl|\frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_k^x}(s') - \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_k^x}(s)\biggr| \leq \lambda_{\mathrm{R}}\mathbb{Q}_k^x\biggl(\mathbb{E}\biggl(\frac{\mathrm{d}}{\mathrm{d}s}_{\big|s=s'} q(sI(o)^{-1}) - \frac{\mathrm{d}}{\mathrm{d}s}_{\big|s=s} q(sI(o)^{-1}) \biggm| X\biggr)\biggr),$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

which is at most $Nw^{-2}|s'-s|$ and, therefore, tends to 0 as k tends to ∞. Thus,

$$\liminf_{k\to\infty} g^k_{\mathbb{Q}^x_k}(s) = \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}^x_*}(s).$$

Using the fact that the function h is lower-semicontinuous, we arrive at

$$\liminf_{k\to\infty} h(f^k_x(s) \mid g^k_{\mathbb{Q}^x_k}(s)) \geq h\biggl(f(x,s) \biggm| \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}^x_*}(s)\biggr),$$

as required.

Finally, let $F \notin AC_0$. First, for any $\varepsilon > 0$, there exists $\mathbb{Q}_{S,x} \in \mathscr{P}_\theta$ such that

$$\begin{aligned}
\tilde{\mathcal{I}}(F) &\geq \sup_{\Xi,S}\biggl[\sum_{x\in\Xi^d} |\Lambda_\Xi(x)| h(\mathbb{Q}_{S,x} \mid \mathbb{P}) \\
&\quad + \sum_{i=1}^r |\Lambda_\Xi(x)| h\biggl(\frac{F(\Lambda_\Xi(x)\times(s_{i-1},s_i])}{|\Lambda_\Xi(x)|} \biggm| \Delta\mu_{\mathbb{Q}_{S,x}}(s_i)\biggr)\biggr] - \varepsilon \\
&\geq \sup_{\Xi,S}\biggl[\sum_{x\in\Xi^d}\sum_{i=1}^r |\Lambda_\Xi(x)| h\biggl(\frac{F(\Lambda_\Xi(x)\times(s_{i-1},s_i])}{|\Lambda_\Xi(x)|} \biggm| \Delta\mu_{\mathbb{Q}_{S,x}}(s_i)\biggr)\biggr] - \varepsilon \\
&= \sup_{\Xi,S}\biggl[\sum_{x\in\Xi^d}\sum_{i=1}^r |\Lambda_\Xi(x)| \sup_\rho\biggl[\rho\frac{F(\Lambda_\Xi(x)\times(s_{i-1},s_i])}{|\Lambda_\Xi(x)|} - (\mathrm{e}^\rho - 1)\Delta\mu_{\mathbb{Q}_{S,x}}(s_i)\biggr]\biggr] - \varepsilon
\end{aligned}$$

using also the Legendre transform of the relative entropy. Furthermore, we have

$$|\mu_{\mathbb{Q}_{S,x}}(s_i) - \mu_{\mathbb{Q}_{S,x}}(s_{i-1})| \leq \lambda_\mathrm{R}|\mathbb{Q}_{S,x}(\Gamma(s_i,o) - \Gamma(s_{i-1},o)| \leq N\lambda_\mathrm{R} w^{-2}|s_i - s_{i-1}|,$$

and, hence, for $\rho \geq 0$,

$$\tilde{\mathcal{I}}(F) \geq \rho \sup_{\Xi,S}\biggl[\sum_{x\in\Xi^d}\sum_{i=1}^r [F(\Lambda_\Xi(x)\times(s_{i-1},s_i]) - (\mathrm{e}^\rho - 1)N\lambda_\mathrm{R}|\Lambda_\Xi(x)||s_i - s_{i-1}|]\biggr] - \varepsilon.$$

If F is not right-continuous, there exists a point (x,s) such that $F(x,s) < \lim_{n\to\infty} F(x + 1/n, s + 1/n) = M$. Consider a sequence of finite partitions $(\Xi_n, S_n)_{n\in\mathbb{N}}$ where the cube $(\prod_{j=1}^d (x_j, x_j + 1/n]) \times (s, s+1/n]$ is contained in (Ξ_n, S_n) for all $n \in \mathbb{N}$. Then

$$\tilde{\mathcal{I}}(F) \geq \rho\biggl[F\biggl(x + \frac{1}{n}, s + \frac{1}{n}\biggr) - F(x,s) - \frac{(\mathrm{e}^\rho - 1)N\lambda_\mathrm{R} 1}{n^{d+1}}\biggr] - \varepsilon$$

and letting $n \to \infty$ gives $\tilde{\mathcal{I}}(F) \geq \rho[M - F(x,s)] - \varepsilon$ which tends to ∞ as $\rho \to \infty$.

If F is right-continuous but $F \notin AC_0$ there exists $\delta > 0$ and a sequence of measurable sets A_k, with $\nu_{d+1}(A_k) \to 0$ and $\mu_F(A_k) \geq \delta$. Using the regularity of the Lebesgue measure there exists a disjoint union of countably many $(d+1)$-dimensional cuboids such that $A_k \subset \bigcup_l q^k_l$ and $\nu_{d+1}(\bigcup_l q^k_l \setminus A_k) < 1/k$. Then, for every $\rho \geq 0$,

$$\begin{aligned}
\tilde{\mathcal{I}}(F) &\geq \rho \sum_{l=1}^\infty F(q^k_l) - (\mathrm{e}^\rho - 1)N\lambda_\mathrm{R}\nu_{d+1}\biggl(\bigcup_l q^k_l\biggr) - \varepsilon \\
&\geq \rho\mu_F(A_k) - (\mathrm{e}^\rho - 1)N\lambda_\mathrm{R}\biggl(\nu_{d+1}(A_k) + \frac{1}{k}\biggr) - \varepsilon.
\end{aligned}$$

Letting k tend to ∞, we have $\tilde{\mathcal{I}}(F) \geq \rho\delta - \varepsilon$ which tends to ∞ as $\rho \to \infty$. □

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

3.3. Contraction principle and identification of rate function

In this section we apply the contraction principle to derive Theorem 2 from Proposition 2. Consider the function $\Psi\colon \mathcal{K} \to \mathcal{M}(\Lambda_1')$ given by

$$F(\cdot) \mapsto F((\cdot \times [0,\infty)) \cap \{(y,s) \in \Lambda_1' \times [0,\infty) \colon s \le |y|^{-\alpha}\}).$$

Then, the random measure $\Psi(|\Lambda_t'|^{-1} Y^{*,t}(\cdot,\cdot))$ is exponentially equivalent to the random measure $|\Lambda_t'|^{-1} Y^t(t^{-\beta}\cdot)$. Moreover, Ψ is continuous when restricted to the subset $L_{\mathrm{inc},0}(\Lambda_1' \times [0,\infty))$ of $L_{\mathrm{inc}}(\Lambda_1' \times [0,\infty))$ consisting of those F with $\mu_F(\partial M) = 0$, where

$$M = \{(y,s) \in \Lambda_1' \times [0,\infty) : s \le |y|^{-\alpha}\}$$

and $\mathcal{K} \subset L_{\mathrm{inc}}(\Lambda_1' \times [0,\infty))$ denotes the family of all $[0,\infty)$-valued, bounded, increasing, and right-continuous functions. Since the Lebesgue measure of ∂M is 0, the rate function from the LDP of Proposition 2 is infinite on the complement of $\mathcal{K}$. Hence, [10, Lemma 4.1.5] shows that the random fields $\{|\Lambda_t'|^{-1} Y^{*,t}(\cdot,\cdot)\}_{t<1}$ also satisfy an LDP on $\mathcal{K}$. Hence, the contraction principle applies and it remains to identify the rate function. That is, we need to show that

$$\inf_{F\in\mathcal{K},\, G(\cdot)=F(\cdot\mathbf{1}_M)} \int_{\Lambda_1'} \inf_{\mathbb{Q}} h(\mathbb{Q} \mid \mathbb{P}) + \int_0^\infty h\bigg(f(y,s) \,\bigg|\, \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}}(s)\bigg)\,\mathrm{d}s\,\mathrm{d}y$$
$$= \int_{\Lambda_1'} \inf_{\mathbb{Q}} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha}))\,\mathrm{d}y,$$

where $f = \partial F/\partial y \partial s$ and $g = \partial G/\partial y$ denote the Radon–Nikodym derivatives of F and G, respectively. Note that if G was not absolutely continuous, then neither could F be, so that the left-hand side would be ∞. We show that the equality arises as a consequence of two inequalities. First, we show that the left-hand side is at least as large as the right-hand side. As in the proof of Proposition 2, an application of Jensen's inequality implies that

$$\int_0^{|y|^{-\alpha}} h\bigg(f(y,s) \,\bigg|\, \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}}(s)\bigg)\,\mathrm{d}s \ge h\bigg(\int_0^{|y|^{-\alpha}} f(y,s)\,\mathrm{d}s \mid \mu_{\mathbb{Q}}(|y|^{-\alpha})\bigg).$$

The right-hand side is equal to $h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha}))$ if $G(\cdot) = F(\cdot\mathbf{1}_M)$.

The other direction is more involved. First, we proceed as in the proof of Proposition 2 and note that the right-hand side can be approximated using a suitable discretization. To be more precise, let $\{\rho(l)\}_{l=1}^{2^{dk}}$ be a subdivision of Λ_1' into congruent cubes of side length $\delta(k) = |\Lambda_1'|^{1/d} 2^{-k}$. The point in the lth cube which minimizes the distance to the origin will be denoted by $y_{k,l}$. In the first step of the discretization, we replace the expression $\mu_{\mathbb{Q}}(|y|^{-\alpha})$ by $\mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})$.

Lemma 9. *We have*

$$\limsup_{k\to\infty} \sum_{l=1}^{2^{dk}} \int_{\rho(l)} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha}))\,\mathrm{d}y$$
$$\le \int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha}))\,\mathrm{d}y.$$

Proof. First, note that for every $l \in \{1, \dots, 2^{dk}\}$, $y \in \rho(l)$, and $\mathbb{Q} \in \mathcal{P}_\theta$, we have

$$\begin{aligned} &h(g(y) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})) - h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha})) \\ &\quad \le g(y) \log \frac{\mu_{\mathbb{Q}}(|y|^{-\alpha})}{\mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})} + |\mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha}) - \mu_{\mathbb{Q}}(|y|^{-\alpha})| \\ &\quad \le |\mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha}) - \mu_{\mathbb{Q}}(|y|^{-\alpha})|, \end{aligned}$$

where the last inequality follows from the choice of $y_{k,l}$. In particular, the right-hand side is always bounded above by 1. Moreover, for $\varepsilon > 0$, we let $A_\varepsilon = \{l \in \{1, \dots, 2^{dk}\} \colon \min_{y \in \rho(l)} |y| < \varepsilon\}$ denote the set of indices of cubes that are close to the origin. Then, the Lipschitz assumption implies that, for every $l \notin A_\varepsilon$, $y \in \rho(l)$ and $\mathbb{Q} \in \mathcal{P}_\theta$,

$$|\mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha}) - \mu_{\mathbb{Q}}(|y|^{-\alpha})| \le N\alpha |y_{k,l}|^{-\alpha-1} |y - y_{k,l}| \le N\alpha \varepsilon^{-\alpha-1} \sqrt{d} \delta(k)^{-1}.$$

Hence, for sufficiently large $k \ge 1$, the difference

$$\begin{aligned} &\sum_{l=1}^{2^{dk}} \int_{\rho(l)} \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})) \,\mathrm{d}y \\ &\qquad - \bigg(\int_{\Lambda_1'} \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha})) \,\mathrm{d}y \bigg) \end{aligned}$$

is bounded above by $|\rho(1)| \# A_\varepsilon + N\alpha\varepsilon^{-\alpha-1}\sqrt{d}\delta(k) \le 2^d \varepsilon^d + N\alpha\varepsilon^{-\alpha-1}\sqrt{d}\delta(k)$, Since $\varepsilon > 0$ was arbitrary, this completes the proof. □

The next lemma is proved similarly to Proposition 2 using Jensen's inequality and a discretization of the integral. We omit the proof.

Lemma 10. *Let $k \ge 1$ and $1 \le l \le 2^{dk}$ be arbitrary. Then,*

$$\begin{aligned} &|\rho(l)|^{-1} \int_{\rho(l)} \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})) \,\mathrm{d}y \\ &\qquad \ge \inf_{\mathbb{Q} \in \mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(|\rho(l)|^{-1} G(\rho(l)) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})) \,\mathrm{d}y. \end{aligned}$$

Now that we have discretized the integral, we can define approximations $F^{(k)}$ to the desired function F. For this purpose, we first need to construct certain minimizers. Recall from Corollary 3 that the function $\mathbb{Q} \mapsto \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha})$ is continuous on every set of the form $\{\mathbb{Q} \colon h(\mathbb{Q} \mid \mathbb{P}) \le K\}$ for some $K < \infty$. Therefore, the function

$$\mathbb{Q} \mapsto h(\mathbb{Q} \mid \mathbb{P}) + h(|\rho(l)|^{-1} G(\rho(l)) \mid \mu_{\mathbb{Q}}(|y_{k,l}|^{-\alpha}))$$

is lower semicontinuous, and we let $\mathbb{Q}_{k,l}$ be one of its minimizers. Now, define measurable functions $f^{(k)} \colon \Lambda_1' \to [0, \infty]$, $k \ge 1$, by

$$f^{(k)}(y, s) = \begin{cases} \dfrac{|\rho(l)|^{-1} G(\rho(l)) (\mathrm{d}/\mathrm{d}s) \mu_{\mathbb{Q}_{k,l}}(s)}{\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})} & \text{if } y \in \rho(l) \text{ and } s \le |y_{k,l}|^{-\alpha}, \\ \dfrac{\mathrm{d}}{\mathrm{d}s} \mu_{\mathbb{Q}_{k,l}}(s) & \text{if } y \in \rho(l) \text{ and } s > |y_{k,l}|^{-\alpha}. \end{cases}$$

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

Here, we make the convention that the first line is equal to 0 if $\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) = G(\rho(l)) = 0$ and is equal to ∞ if $\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) = 0$, but $G(\rho(l)) \neq 0$. Furthermore, we let $F^{(k)}$ denote the distribution function of the measure with density $f^{(k)}(y, s)$. Then, for every $y \in \rho(l)$, the expression $\inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \int_0^\infty h(f^{(k)}(y, s) \mid (\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}}(s))\,\mathrm{d}s$ is bounded above by

$$h(\mathbb{Q}_{k,l} \mid \mathbb{P}) + \int_0^{|y_{k,l}|^{-\alpha}} h\bigg(\frac{|\rho(l)|^{-1}G(\rho(l))(\mathrm{d}/\mathrm{d}s)\mu_{\mathbb{Q}_{k,l}}(s)}{\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})} \,\bigg|\, \frac{\mathrm{d}}{\mathrm{d}s}\mu_{\mathbb{Q}_{k,l}}(s)\bigg)\,\mathrm{d}s$$
$$= h(\mathbb{Q}_{k,l} \mid \mathbb{P}) + h(|\rho(l)|^{-1}G(\rho(l)) \mid \mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})).$$

From the favorable rate function in Proposition 2, the functions $(F^{(k)})_{k\geq 1}$ have an accumulation point, and from lower-semicontinuity, therefore, we have

$$\int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \int_0^\infty h(f(y, s) \mid \mu_{\mathbb{Q}}(s))\,\mathrm{d}s$$
$$\leq \int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + h(g(y) \mid \mu_{\mathbb{Q}}(|y|^{-\alpha}))\,\mathrm{d}y.$$

Hence, it remains to show that the measures induced by $G(\cdot)$ and $F(\cdot\mathbf{1}_M)$ coincide. In order to prove this claim, we first show that $F^{(k)}(M^{(k)} \setminus M)$ tends to 0 as k tends to ∞, where $M^{(k)} = \{(y, s) \in \Lambda_1' \times [0, \infty)\colon y \in \rho(l) \text{ and } s \leq |y_{k,l}|^{-\alpha}\}$.

Lemma 11. *The expression $F^{(k)}(M^{(k)} \setminus M)$ tends to 0 as k tends to ∞.*

Proof. First, observe that $F^{(k)}(M^{(k)} \setminus M)$ can be expressed as

$$F^{(k)}(M^{(k)} \setminus M) = \sum_{l=1}^{2^{dk}} G(\rho(l))|\rho(l)|^{-1} \int_{\rho(l)} \frac{\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) - \mu_{\mathbb{Q}_{k,l}}(|y|^{-\alpha})}{\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})}\,\mathrm{d}y.$$

Now, for $\varepsilon > 0$, introduce the set

$$A_\varepsilon = \{l \in \{1, \dots, 2^{dk}\}\colon \min_{y\in\rho(l)} |y| < \varepsilon \text{ or } \max_{y\in\rho(l)} |y|^{-\alpha} > ws_{\min} - \varepsilon\}$$

of indices whose associated cubes are far away from the origin and the boundary of the ball $B_{(ws_{\min})^{-1/\alpha}}(o)$. Hence, we arrive at

$$F^{(k)}(M^{(k)} \setminus M) = \alpha\varepsilon^{-\alpha-1}N\sqrt{d}\delta(k) \sum_{l\notin A_\varepsilon} \frac{G(\rho(l))}{\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})} + r(\varepsilon),$$

where $r(\varepsilon)$ tends to 0 as ε tends to 0. Since the sum above consists of at most 2^{dk} summands, it is enough to consider those $l \notin A_\varepsilon$ that satisfy $|\rho(l)|^{-1}G(\rho(l)) > \mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})$. Now, note that if $l \notin A_\varepsilon$ then $\mu_{\mathbb{P}}(|y_{k,l}|^{-\alpha}) \geq 1/K$ for some sufficiently large $K = K(\varepsilon)$ not depending on k and l. We also claim that $\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) \geq 1/K$. Once this is shown, the proof is complete.

Suppose that $\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) < 1/K$. First, if $|\rho(l)|^{-1}G(\rho(l)) \geq \mu_{\mathbb{P}}(|y_{k,l}|^{-\alpha})$ then

$$h(|\rho(l)|^{-1}G(\rho(l)) \mid \mu_{\mathbb{P}}(|y_{k,l}|^{-\alpha})) < h(|\rho(l)|^{-1}G(\rho(l)) \mid \mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})),$$

which contradicts the minimality of $\mathbb{Q}_{k,l}$. Otherwise, set $\mathbb{Q}^* = \lambda\mathbb{P} + (1 - \lambda)\mathbb{Q}_{k,l}$, where $\lambda \in [0, 1)$ is chosen such that

$$\lambda\mu_{\mathbb{P}}(|y_{k,l}|^{-\alpha}) + (1 - \lambda)\mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha}) = |\rho(l)|^{-1}G(\rho(l)).$$

Then, since the specific relative entropy h (introduced in Section 1) is an affine function,

$$h(\mathbb{Q}^* \mid \mathbb{P}) + h(|\rho(l)|^{-1}G(\rho(l)) \mid \mu_{\mathbb{Q}^*}(|y_{k,l}|^{-\alpha})) = (1-\lambda)h(\mathbb{Q}_{k,l} \mid \mathbb{P}).$$

As the right-hand side is strictly smaller than

$$h(\mathbb{Q}_{k,l} \mid \mathbb{P}) + h(|\rho(l)|^{-1}G(\rho(l)) \mid \mu_{\mathbb{Q}_{k,l}}(|y_{k,l}|^{-\alpha})),$$

this contradicts again the minimality of $\mathbb{Q}_{k,l}$. □

Now, we can complete the proof of Theorem 2.

Proof of Theorem 2. By Lemmas 9 and 10, it suffices to show that $G(f) = F(f\mathbf{1}_M)$ holds for any $f\colon \Lambda_1' \to [0,\infty)$ of the form $f = \mathbf{1}_{\rho(l_0)}$ for some $1 \le l_0 \le 2^{dk_0}$. Now, we can argue as follows:

$$|F(f\mathbf{1}_M) - G(f)| \le \limsup_{k\to\infty} |F^{(k)}(f\mathbf{1}_{M^{(k)}}) - G(f)| + \limsup_{k\to\infty} F^{(k)}(f\mathbf{1}_{M^{(k)}\setminus M}).$$

from Lemma 11 we see that the second summand is 0. Moreover, from the definition of $F^{(k)}$, we see that $F^{(k)}(f\mathbf{1}_{M^{(k)}}) = G^{(k)}(f)$, where $G^{(k)}$, $k \ge k_0$, denotes the measure with locally constant density $g^{(k)}(y) = |\rho(l)|^{-1}G(\rho(l))$. Hence,

$$G^{(k)}(f) = G^{(k)}(\rho(l_0)) = \sum_{\rho(l)\subset\rho(l_0)} G(\rho(l)) = G(\rho(l_0)) = G(f),$$

as required. □

3.4. Proof of Corollary 1

Proof. The upper estimate is a direct consequence of the upper bound in Theorem 2,

$$\begin{aligned}\limsup_{t\to 0} |\Lambda_t'|^{-1}\log p_t &= \limsup_{t\to 0} |\Lambda_t'|^{-1}\log \mathbb{P}(|\Lambda_t'|^{-1}Y^t(t^{-\beta}\cdot) = 0)\\ &\le -\int_{\Lambda_1'} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q}\mid\mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|y|^{-\alpha}, o)))\,\mathrm{d}y.\end{aligned}$$

For the lower estimate, first note that $p_t = \mathbb{E}\exp(-\lambda_{\mathrm{R}}\int_{\Lambda_t'}\Gamma(t^{-1}\ell(y), y)\,\mathrm{d}y)$, where $\Gamma(a,y) = \mathbb{E}(q(aI(y)^{-1}) \mid X)$ is a nonlocal function of the transmitter process. In order to be able to apply [16, Theorem 3.1], we need to establish a translation-invariant setting using discretization of the integrand. To be more precise, we subdivide Λ_t' into 2^{dn} subcubes Λ_t^i of side length $2^{1-n}|ws_{\min}t|^{-\beta}$ and let y_i denote the corresponding element of the subcube Λ_t^i which is closest to the origin. Then,

$$p_t \ge \mathbb{E}\exp\biggl(-\lambda_{\mathrm{R}}\sum_{i=1}^{2^{dn}}\int_{\Lambda_t^i}\Gamma(t^{-1}\ell(|y_i|), y)\,\mathrm{d}y\biggr) \ge \mathbb{E}\exp\biggl(-\lambda_{\mathrm{R}}\sum_{i=1}^{2^{dn}}\int_{\Lambda_t^i}\Gamma^b(t^{-1}\ell(|y_i|), y)\,\mathrm{d}y\biggr),$$

where $\Gamma^b(a,y) = \mathbb{E}(q(aI^b(y)^{-1}) \mid X)$. Furthermore, let $X^{\mathrm{per},i}$ be the configuration obtained after extending the configuration of the marked Poisson point process X in the subcube Λ_t^i periodically in the entire Euclidean space $\mathbb{R}^d$. The error made in replacing $\sum_{i=1}^{2^{dn}}\int_{\Lambda_t^i}\Gamma^b(t^{-1}\ell(y_i), y)$

by $\sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)$ is negligible in the large deviation principle where $\Gamma^{\mathrm{per},b}(a, y) = \mathbb{E}(q(aI^b(y)^{-1}) \mid X^{\mathrm{per},i})$, indeed

$$
\begin{aligned}
\int_{\Lambda_t^i} &[\Gamma^b(t^{-1}\ell(y_i), y) - \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)]\,\mathrm{d}y \\
&\le N(w^2 t)^{-1} \int_{\Lambda_t^i} [I^b(o, y, X) - I^b(o, y, X^{\mathrm{per},i})]\,\mathrm{d}y \\
&\le N(s_{\min}^2 w^2 t)^{-1}\bigg[\sum_{X_j \in X,\, X_j \notin \Lambda_t^i,\, X_j \in \Lambda_{t,b}^i} \int_{\Lambda_t^i} \ell_b(|X_j - y|)\,\mathrm{d}y \\
&\qquad + \sum_{X_j \in X^{\mathrm{per},i},\, X_j \notin \Lambda_t^i,\, X_j \in \Lambda_{t,b}^i} \int_{\Lambda_t^i} \ell_b(|X_j - y|)\,\mathrm{d}y\bigg] \\
&\le N(s_{\min}^2 w^2 t)^{-1}[X(\Lambda_{t,b}^i \setminus \Lambda_t^i) + X^{\mathrm{per},i}(\Lambda_{t,b}^i \setminus \Lambda_t^i)] \int_{B_b(o)} \ell_b(|y|)\,\mathrm{d}y, \qquad (4)
\end{aligned}
$$

where $\Lambda_{t,b}^i$ denotes the volume Λ_t^i joined with its b-boundary. Hence, for all $\varepsilon > 0$,

$$
\begin{aligned}
\mathbb{E}\exp&\bigg(-\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^b(t^{-1}\ell(|y_i|), y)\,\mathrm{d}y\bigg) \\
&\ge \mathrm{e}^{-\varepsilon|\Lambda_t'|}\mathbb{E}\exp\bigg(-\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^{\mathrm{per},b}(t^{-1}\ell(|y_i|), y)\,\mathrm{d}y\bigg) \\
&\quad - \mathbb{P}\bigg(\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^b(t^{-1}\ell(y_i), y) - \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)\,\mathrm{d}y \ge \varepsilon|\Lambda_t'|\bigg).
\end{aligned}
$$

By (4), the second line is bounded from below by

$$
-2\mathbb{P}\bigg(\lambda_{\mathrm{R}} 2N(s_{\min}^2 w^2 t)^{-1} \int_{B_b(o)} \ell_b(|y|)\,\mathrm{d}y\, X(\Lambda_{t,b}^1 \setminus \Lambda_t^1) \ge \varepsilon|\Lambda_t^1|\bigg).
$$

But this goes to 0 on an exponential scale infinitely fast by Lemma 1. Hence,

$$
\begin{aligned}
\lim_{t\to 0} &|\Lambda_t'|^{-1} \log \mathbb{E}\exp\bigg(-\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^b(t^{-1}\ell(y_i), y)\,\mathrm{d}y\bigg) \\
&\ge \lim_{t\to 0} |\Lambda_t'|^{-1} \log \mathbb{E}\exp\bigg(-\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)\,\mathrm{d}y\bigg) \\
&= 2^{-dn} \sum_{i=1}^{2^{dn}} \lim_{t\to 0} |\Lambda_t^i|^{-1} \log \mathbb{E}\exp\bigg(-\lambda_{\mathrm{R}} \int_{\Lambda_t^i} \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)\,\mathrm{d}y\bigg),
\end{aligned}
$$

where we used the independence of the $X^{\mathrm{per},i}$ with respect to i in the second line. Now we are in the position to apply [16, Theorem 3.1] and write

$$\lim_{t\to 0} |\Lambda_t'|^{-1} \log \mathbb{E} \exp\bigg(-\lambda_{\mathrm{R}} \sum_{i=1}^{2^{dn}} \int_{\Lambda_t^i} \Gamma^{\mathrm{per},b}(t^{-1}\ell(y_i), y)\,\mathrm{d}y\bigg)$$
$$\geq -2^{-dn} \sum_{i=1}^{2^{dn}} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma^b(|y_i|^{-\alpha}, o))),$$

using the continuity of $\Gamma^{\mathrm{per},b}$ ensured by the truncation of the interference. Note that

$$\limsup_{b\to\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma^b(|y_i|^{-\alpha}, o))) \leq \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|y_i|^{-\alpha}, o))).$$

Indeed, let $\mathbb{Q}_0$ be a minimizer of the right-hand side, then

$$\limsup_{b\to\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma^b(|y_i|^{-\alpha}, o))) \leq h(\mathbb{Q}_0 \mid \mathbb{P}) + \lambda_{\mathrm{R}} \limsup_{b\to\infty} \mathbb{Q}_0(\Gamma^b(|y_i|^{-\alpha}, o))$$

and it suffices to show that

$$|\mathbb{Q}_0(\Gamma^b(|y|^{-\alpha}, o)) - \mathbb{Q}_0(\Gamma(|y|^{-\alpha}, o))| \leq N|y|^{-\alpha} w^{-2} \mathbb{Q}_0\bigg(\sum_{X_j \not\subset \Lambda_b} \ell(|X_j|)\bigg)$$

tends to 0 as b tends to ∞. But this holds since $\mathbb{Q}_0$ is a translation-invariant point process. In order to perform the large-n limit, we have to show that

$$y \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta} (h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|y|^{-\alpha}, o)))$$

is continuous. But this also holds since

$$\Big|\inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|y|^{-\alpha}, o))) - \inf_{\mathbb{Q}\in\mathcal{P}_\theta} h(\mathbb{Q} \mid \mathbb{P}) + \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|x|^{-\alpha}, o)))\Big|$$
$$\leq \sup_{\mathbb{Q}\in\mathcal{P}_\theta} |\lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|y|^{-\alpha}, o)) - \lambda_{\mathrm{R}}\mathbb{Q}(\Gamma(|x|^{-\alpha}, o))|$$
$$\leq N w^{-1} \lambda_{\mathrm{R}} ||y|^{-\alpha} - |x|^{-\alpha}|. \qquad \square$$

4. Importance sampling

In this section we show how the LDPs derived in Theorems 1 and 2 can be used to devise an importance sampling scheme improving the accuracy of basic Monte Carlo approaches for estimating the probability of observing unlikely configurations of connectable receivers. Theorems 1 and 2 imply that such probabilities generally tend to 0 exponentially quickly, so that basic Monte Carlo estimators perform poorly.

The general heuristic for devising importance sampling schemes is the following. Instead of sampling the transmitters according to their true distribution, the simulation is performed by using a modified law under which the considered rare event is more likely. An appropriate reweighting using likelihood ratios ensures the unbiasedness of the new estimator. For a more detailed discussion of the general technique of importance sampling, we refer the reader to [2] and [19].

https://doi.org/10.1017/apr.2016.65 Published online by Cambridge University Press

In principle, Theorems 1 and 2 provide precise descriptions of the asymptotically exponentially optimal change of measure, in the sense that the modified law of transmitters should be given by suitable Gibbs point processes. However, as these distributions just arise as minimizers of fairly complicated functionals, it is difficult to use them for computational purposes. Still, by performing this minimization in the restricted class of Poisson point processes, we can achieve substantial accuracy benefits.

We only provide a proof of concept for the use of importance sampling, and, therefore, assume a specific parameter constellation in the following. First, we fix $d = 2$, $w = \lambda_{\mathrm{R}} = \lambda_{\mathrm{T}} = 1$, and assume that the path-loss function is given by $\ell(r) = r^{-4}$. Moreover, we assume that there is no random environment, and that transmission powers and fading random variables are constant and equal to 1. Note that this choice is not covered by the assumptions for Theorems 1 and 2. Nevertheless, our simulation results illustrate that variance reduction through importance sampling also hold under weaker conditions than the ones assumed in Theorems 1 and 2.

4.1. Importance sampling related to Theorem 1

Since we have assumed that there is no random environment and that transmission powers and fading variables are constant, the minimization in the rate function of Theorem 1 is performed only over stationary point processes of transmitters and receivers. As mentioned above, this minimization is intractable in its full generality. Nevertheless, in this section, we show that if minimization is performed only in the class of Poisson point processes, then the problem becomes tractable. In fact, we provide an example problem, where the minimization can be reduced to a standard two-dimensional constrained minimization problem, where the constraint is given in terms of certain special functions. The disadvantage of this approach is that solving the minimization problem in a restricted class of point process will not automatically lead to good choices for the importance sampling. This will become apparent from the simulation results discussed below.

We assume that $t = 1$ and consider events of the form

$$A_{n,a} = \left\{ \frac{1}{|\Lambda_n|} \sum_{X_i \in \Lambda_n} \#Y^{(i)} < a \right\},$$

i.e. the event that the (spatially) averaged number of connectable receivers associated with transmitters in the cube Λ_n is less than a. Now, we explain how to implement an importance sampling scheme based on the LDP. A related importance sampling scheme for a Poisson point process on the real line has already been considered in [17], but for the convenience of the reader, we present some details in our situation.

In order to estimate the probability of the event $A_{n,a}$, we simulate the Poisson point processes with new intensities $\mu_{\mathrm{R}} > 0$ and $\mu_{\mathrm{T}} > 0$ in Λ_n. Then, the likelihood ratio of a Poisson point process with intensity 1 with respect to these point processes is given by

$$\exp(|\Lambda_n|(\mu_{\mathrm{R}} - 1) + |\Lambda_n|(\mu_{\mathrm{T}} - 1))\mu_{\mathrm{R}}^{-X(\Lambda_n)}\mu_{\mathrm{T}}^{-Y(\Lambda_n)}.$$

Hence, an unbiased estimator for $\mathbb{P}(A_{n,a})$ is given by

$$\widehat{p}_{n,a,\mu_{\mathrm{T}},\mu_{\mathrm{R}}} = \exp(|\Lambda_n|(\mu_{\mathrm{R}} - 1) + |\Lambda_n|(\mu_{\mathrm{T}} - 1))\mu_{\mathrm{R}}^{-X(\Lambda_n)}\mu_{\mathrm{T}}^{-Y(\Lambda_n)}\mathbf{1}_{A_{n,a}}.$$

In order to take into account edge effects, we also generate transmitters and receivers with the unmodified intensity in a small environment around Λ_n. We can obtain estimates $\widehat{p}$ and $\widehat{v}$ of

TABLE 1: Comparison of the simulation results for the expectation and variance of the considered importance sampling estimators with transmitter and receiver intensities $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}})$.

$(\mu_{\mathrm{R}}, \mu_{\mathrm{T}})$	$\widehat{p}$	$\widehat{v}$
(1, 1)	3.31×10^{-4}	3.31×10^{-4}
(0.832, 0.984)	3.12×10^{-4}	3.69×10^{-4}
(0.892, 0.989)	3.29×10^{-4}	5.10×10^{-5}

the expectation and variance of $\widehat{p}_{n,a,\mu_{\mathrm{T}},\mu_{\mathrm{R}}}$ by considering the sample average and variance of $N \geq 1$ i.i.d. copies generated using Monte Carlo simulation.

This leaves the question as of how to find good choices for μ_{R} and μ_{T}. In a first attempt, choose these parameters to minimize the large deviation rate function appearing in Theorem 1. If $\mathbb{Q} \in \mathcal{P}_\theta$ is the distribution of independent Poisson point processes of receivers and transmitters with intensities μ_{R} and μ_{T}, then the relative entropy $h(\mathbb{Q} \mid \mathbb{P})$ is given by $(\mu_{\mathrm{R}} \log \mu_{\mathrm{R}} - \mu_{\mathrm{R}} + 1) + (\mu_{\mathrm{T}} \log \mu_{\mathrm{T}} - \mu_{\mathrm{T}} + 1)$. Hence, to determine the optimal intensities $(\lambda_{\mathrm{R,opt}}, \lambda_{\mathrm{T,opt}})$ according to Theorem 1, we need to minimize this formula under the constraint $\mathbb{Q}^*(\#Y^{(o)}) < a$. Next, we express this constraint in terms of certain special functions. First, by Campbell's theorem,

$$\begin{aligned}\mathbb{Q}^*(\#Y^{(o)}) &= \mu_{\mathrm{T}}\mu_{\mathrm{R}} \int_{B_1(o)} \mathbb{Q}^*\bigg(\frac{|x|^{-4}}{1 + \sum_{i \geq 1} |X_i|^{-4}} \geq 1\bigg)\,\mathrm{d}x \\ &= \mu_{\mathrm{T}}\mu_{\mathrm{R}} 2\pi \int_0^1 r\mathbb{P}\bigg(\sum_{i \geq 1} |X_i|^{-4} \leq \mu_{\mathrm{T}}^2(r^4 - 1)\bigg)\,\mathrm{d}r,\end{aligned}$$

where in the last line we used that scaling by $1/\sqrt{\mu_{\mathrm{T}}}$ transforms a Poisson point process with intensity 1 to a Poisson point process with intensity μ_{T}. Moreover, $\sum_{i \geq 1} |X_i|^{-4}$ is distributed according to an inverse gamma distribution with parameters $\frac{1}{2}$ and $\pi^3/4$. In particular, $\mathbb{P}(\sum_{i \geq 1} |X_i|^{-4} \leq s) = \pi^{-1/2}\gamma(\frac{1}{2}, -\pi^3/4s)$, where $\gamma(\cdot,\cdot)$ denotes the incomplete gamma function. Now it is easy to check that

$$\mathbb{Q}^*(\#Y^{(o)}) = \mu_{\mathrm{R}}\mu_{\mathrm{T}} 2\pi \int_0^1 r\gamma\bigg(\frac{1}{2}\bigg)\bigg(\frac{\pi^3}{4\mu_{\mathrm{T}}^2(r^4 - 1)}\bigg)\,\mathrm{d}r = \mu_{\mathrm{R}}\mu_{\mathrm{T}}\pi \exp\bigg(\frac{\pi^3}{4\mu_{\mathrm{T}}^2}\bigg)\mathrm{erfc}\bigg(\frac{\pi^{3/2}}{2\mu_{\mathrm{T}}}\bigg),$$

where 'erfc' denotes the complimentary error function. For instance, if we choose $a = \frac{1}{2}$ then $(\lambda_{\mathrm{R,opt}}, \lambda_{\mathrm{T,opt}}) \approx (0.832, 0.984)$.

In order to assess the accuracy improvements that can be achieved with this importance sampling scheme, we performed a Monte Carlo analysis. We fixed $a = \frac{1}{2}$, $n = 25$, and performed $N = 10\,000\,00$ simulation runs. We consider three different parameter choices for the importance sampling intensities $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}})$. First, we consider the case of basic Monte Carlo simulation; that is, $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}}) = (1, 1)$. Second, we take the intensities that are obtained from the large deviation analysis performed above, i.e. $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}}) = (0.832, 0.984)$. Third, we estimate $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}})$ from a simple cross-entropy scheme. That is, we performed a pilot run of 100 000 basic Monte Carlo simulations and determined the average intensities under the condition that the rare event occurs. We obtain $(\mu_{\mathrm{R}}, \mu_{\mathrm{T}}) = (0.892, 0.989)$. We refer the reader to [19] for details on the general cross-entropy technique. The results for $\widehat{p}$ and $\widehat{v}$ are reported in Table 1.

In particular, we would like to draw the attention to an important observation: the estimator that is obtained as the solution of the optimization based on our large deviation principle actually has a higher variance than the basic Monte Carlo estimator. Given the close relation between the large deviation theory and asymptotically optimal change of measures, this might come as a surprise at first sight. However, since we performed our optimization not in the full class of stationary point processes, but only considered Poisson point processes, the simulation output does not contradict this intuition. In fact, considered from a different perspective, the simulation results provide evidence that the optimal change of measure is rather far (in the Kullback–Leibler distance) from being a Poisson point process. In contrast, performing the change of measure with the intensities obtained from the pilot run shows that for the considered example, a more than seven-fold variance reduction can be achieved.

The discussion in the previous paragraph raises the legitimate question as to whether the change of measures deduced from the large deviation result are of any practical use for importance sampling. Indeed, in the example described above, the intensities that lead to the seven-fold decrease in variance could be found without reference to the LDP; namely, by an 'educated guess' (or rather 'cross-entropy'). Nevertheless, when considering importance sampling in the setting of Corollary 1, finding a good importance sampling change of measure would involve 'guessing' a continuous family of parameters, which is substantially more involved than what we have performed above. In contrast, a simple analysis of the large deviation rate function provides immediately a useful heuristic for the shape of the curve.

4.2. Importance sampling related to Theorem 2

Finally, we investigate importance sampling techniques related to Theorem 2. We consider the specific setting of Corollary 1, i.e. estimation of the isolation probability p_t for small values of t. Similar to the situation considered in Section 4.1, the full minimization problem is intractable, so that we restrict our attention to the class of homogeneous Poisson point processes. However, the situation is slightly different from the one considered in Section 4.1. Instead of globally optimizing a transmitter and receiver intensity, we now have the freedom to choose a different intensity for each point in Λ_1'. Due to isotropy, this reduces to the task of choosing an optimal intensity $\lambda_{\mathrm{opt}}(r)$ for each $r \in [0, 1]$. This optimal intensity must minimize the following expression that can be derived from the variational characterization in Corollary 1:

$$\lambda_{\mathrm{opt}}(r)\log\lambda_{\mathrm{opt}}(r) - \lambda_{\mathrm{opt}}(r) + 1 + \mathbb{P}\bigg(r^{-4} \geq 1 + \lambda_{\mathrm{opt}}(r)^2 \sum_{i\geq 1} |X_i|^{-4}\bigg).$$

This is a standard minimization problem that can be solved by finding the roots of the derivative with respect to $\lambda_{\mathrm{opt}}(r)$. After some simplifications, we arrive at

$$\log\lambda_{\mathrm{opt}}(r) = \frac{\pi}{\sqrt{r^{-4}-1}}\exp\bigg(\frac{\pi^3\lambda_{\mathrm{opt}}(r)^2}{4(r^{-4}-1)}\bigg).$$

This equation can be solved numerically; a plot of this solution is shown in Figure 2.

As in the previous example, in order to assess the actual accuracy improvements for the estimation of p_t that can be achieved with this importance sampling scheme, we performed a prototypical Monte Carlo analysis. We fixed $t = 0.002$ and performed $N = 1\,000\,000$ simulation runs. See Table 2.

In contrast to the previous example, we see that the importance sampling estimator derived from large deviation theory provides substantial benefits. Indeed, the variance is reduced by approximately 78%. Furthermore, applying the cross-entropy technique for the present example would be substantially more involved than in the previous example. Indeed, instead of simply

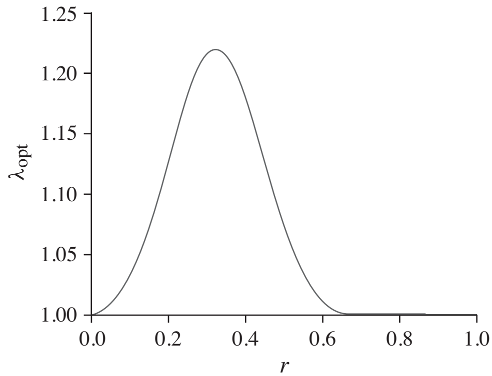

FIGURE 2: Plot of the optimal density $\lambda_{\text{opt}}(r)$ at distance r from the origin.

TABLE 2: Comparison of the simulation results for the expectation and variance of the basic versus the importance sampling estimator.

	$\widehat{p}$	$\widehat{v}$
$\lambda(\cdot) \equiv 1$	7.72×10^{-6}	8.22×10^{-10}
$\lambda(\cdot) \equiv \lambda_{\text{opt}}(\cdot)$	7.70×10^{-6}	1.78×10^{-10}

estimating two parameters, we would need to extract an entire curve from the pilot runs, so that proper statistical tools would be needed to estimate such a functional object from the data.

Acknowledgements

The authors thank W. König for interesting discussions, and for proposing the topic of this paper and the use of stationary empirical fields. The authors thank M. Renger for the hint to apply the contraction principle for deriving Theorem 2. We also thank T. Brereton for pointing out the reference [17]. This research was supported by the Leibniz group on Probabilistic Methods for Mobile Ad-Hoc Networks.

References

[1] ADAMS, S., COLLEVECCHIO, A. AND KÖNIG, W. (2011). A variational formula for the free energy of an interacting many-particle system. *Ann. Prob.* **39,** 683–728.

[2] ASMUSSEN, S. AND GLYNN, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.

[3] BACCELLI, F. AND BŁASZCZYSZYN, B. (2009). Stochastic geometry and wireless networks, Vol. 1, Theory. *Foundations Trends Networking* **3,** 249–449.

[4] BACCELLI, F. AND BŁASZCZYSZYN, B. (2009). Stochastic geometry and wireless networks, Vol. 2, Applications. *Foundations Trends Networking* **4,** 1–312.

[5] BACCELLI, F. AND ZHANG, X. (2015). A correlated shadowing model for urban wireless networks. In *2015 IEEE Conference of Computer Communications (INFOCOM)*, IEEE, New York, pp. 801–809.

[6] BALDI, P., FRIGESSI, A. AND PICCIONI, M. (1993). Importance sampling for Gibbs random fields. *Ann. Appl. Prob.* **3,** 914–933.

[7] BANGERTER, B., TALWAR, S., AREFI, R. AND STEWART, K. (2014). Networks and devices for the 5G era. *IEEE Commun. Magazine* **52,** 90–96.

[8] BOUCHERON, S., LUGOSI, G. AND MASSART, P. (2013). *Concentration Inequalities*. Oxford University Press.

[9] DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*, Vol. I, 2nd edn. Springer, New York.

[10] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.
[11] EICHELSBACHER, P. AND SCHMOCK, U. (1998). Exponential approximations in completely regular topological spaces and extensions of Sanov's theorem. *Stoch. Process. Appl.* **77,** 233–251.
[12] GANESH, A. J. AND TORRISI, G. L. (2008). Large deviations of the interference in a wireless communication model. *IEEE Trans. Inf. Theory* **54,** 3505–3517.
[13] GEORGII, H.-O. (1993). Large deviations and maximum entropy principle for interacting random fields on $\mathbf{Z}^d$. *Ann. Prob.* **21,** 1845–1875.
[14] GEORGII, H.-O. (1994). Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction. *Prob. Theory Relat. Fields* **99,** 171–195.
[15] GEORGII, H.-O. (2011). *Gibbs Measures and Phase Transitions*, 2nd edn. De Gruyter, Berlin.
[16] GEORGII, H.-O. AND ZESSIN, H. (1993). Large deviations and the maximum entropy principle for marked point random fields. *Prob. Theory Relat. Fields* **96,** 177–204.
[17] GIESECKE, K. AND SHKOLNIK, A. (2014). Optimal importance sampling of default losses. Preprint. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1924135.
[18] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.
[19] KROESE, D. P., TAIMRE, T. AND BOTEV, Z. I. (2011). *Handbook of Monte Carlo Methods*. John Wiley, Hoboken, NJ.
[20] MECKE, J. (1967). Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. *Z. Wahrscheinlichkeitsth.* **9,** 36–58.
[21] PENROSE, M. (2003). *Random Geometric Graphs*. Oxford University Press.
[22] SCHREIBER, T. AND YUKICH, J. E. (2005). Large deviations for functionals of spatial point processes with applications to random packing and spatial graphs. *Stoch. Process. Appl.* **115,** 1332–1356.
[23] SEPPÄLÄINEN, T. AND YUKICH, J. E. (2001). Large deviation principles for Euclidean functionals and other nearly additive processes. *Prob. Theory Relat. Fields* **120,** 309–345.
[24] TORRISI, G. L. AND LEONARDI, E. (2013). Simulating the tail of the interference in a Poisson network model. *IEEE Trans. Inf. Theory* **59,** 1773–1787.
[25] TORRISI, G. L. AND LEONARDI, E. (2014). Large deviations of the interference in the Ginibre network model. *Stoch. Systems* **4,** 173–205.