Spatial Birth-Death Wireless Networks

Abishek Sankararaman and François Baccelli

Abstract

We propose and study a novel continuous space-time model for wireless networks which takes into account the stochastic interactions in both space through interference and in time due to randomness in traffic. Our model consists of an interacting particle birth-death dynamics incorporating information-theoretic spectrum-sharing. Roughly speaking, particles (or more generally wireless links) arrive according to a Poisson Point Process on space-time, and stay for a duration governed by the local configuration of points present and then exit the network after completion of a file transfer. We analyze this particle dynamics to derive an explicit condition for time ergodicity (i.e. stability) which is tight. We also prove that when the dynamics is ergodic, the steady-state point process of links (or particles) exhibits a form statistical *clustering*. Based on the clustering, we propose a conjecture which we leverage to derive approximations, bounds and asymptotics on performance characteristics such as delay and mean number of links per unit-space in the stationary regime. The mathematical analysis is combined with discrete event simulation to study the performance of this type of networks.

I. Introduction

We consider the problem of studying the spatial dynamics of Device-to-Device (D2D) or ad-hoc wireless networks. Such wireless networks have received a tremendous amount of attention, due on the one hand to their increasing ubiquity in modern technology and on the other hand to the mathematical challenge in their modeling and performance assessment. Wireless is a broadcast medium and hence the nodes sharing a common spectrum in space interact through the interference they cause to one another. Understanding the limitations due to interference and theoretically optimal protocols in such a static spatial setting has long been considered in network information theory under the interference channel [1]. The full characterization of the interference channel is however a long standing open-problem in network information theory.

In recent years, Stochastic Geometry ([2], [3]) has emerged as a way of assessing performance of links in large-scale networks interacting through interference in space. These tools have been very popular to model and analyze wireless system performance for a variety of network architectures including D2D networks, mobile-ad hoc networks [4] and cellular networks [5]. However, the main drawback in these models is that they do not have a notion of temporal interaction and do not allow one to represent random traffic (they usually rely on a "full-buffer" assumption, i.e., every link always has a packet to transmit).

This additional dimension of interaction among wireless links sharing a common spectrum adds to the complexity of their performance analysis but nonetheless is very crucial to understand network performance. Most prior work aiming at studying the temporal interaction of links model spatial interactions through binary on-off behavior encoded by interference or conflict graphs. The temporal interactions are then modeled using queuing theoretic ideas of flow based models (for ex: [6], [7], [8]). Such flow models have a long history in applied mathematics and engineering. They were initially proposed to study dynamic resource allocation in wired networks ([9], [10]), and were subsequently used to model and study wireless networks. Flow based queuing models have inspired many seminal results in applied probability and networks in the past. The main drawback in employing such models in a wireless scenario however is that the spatial and information-theoretic interactions are overly simplified and not captured precisely.

Motivated by this, we propose a new spatial flow model, which uses the continuum space to model link interaction through interference as prescribed by the information-theoretic setting, and also takes into account the interaction of links across time due to traffic variations. To the best of our knowledge, the analysis of such continuum space-time models for wireless networks has not been considered so far.

The mathematical framework we follow for spatial birth-death process has been studied in different contexts in the probability literature starting from the work of Preston [11]. In recent years, [12] and [13] have also studied in great detail, the problem of general spatial birth and death process which is the basis of our modeling. From a methodological point of view, the work of [14] is the closest in spirit to our work as it also studies a space-time interacting particle process (of a wireline peer-to-peer network). There are several fundamental differences between the model of [14], which is intrinsically stable, and exhibits repulsion, and our model, which is potentially unstable and which exhibits attraction (clustering). Another difference from [14] is that the death-rate (defined later) is a linear-function of the state whereas our model is non-linear because of the information-theoretic formulation, thereby making the analysis more challenging. Nevertheless, we use some of the ideas developed in that paper.

From an engineering viewpoint, this work is motivated by emerging interest in applications like Device-to-Device (D2D) networks and Internet of Things (IoT). These two applications can be viewed as an instance of our abstract mathematical model which is more general. D2D is being considered as a viable networking architecture in future cellular standards to improve system capacity by offloading some traffic from base-station to other mobile devices that have the same content. Some of the more important use cases for such offloading are in a crowded setting (like a stadium or a concert) where there is a huge density of mobile devices. Another important application of D2D is in enabling cellular operators to provide "proximity based services". In such settings, a mobile may access content (which we model as files) from nearby mobile users possessing the content (which may be likely owing to geographical and temporal proximity) rather than from a base-station. Such networking architectures are being envisioned to both reduce the load on the base-stations and also to develop new markets for mobile services. Thus, a snapshot of a D2D network will resemble our model with some mobile devices connecting to and downloading files from other mobile devices that are nearby. IoT is another technology gaining momentum due to the vast market opportunities to develop user applications that leverage the IoT network (for instance in tracking sensors for health, security etc). This network also resembles a wireless ad-hoc network with different things communicating occasionally data to each other or to a central access point using the shared wireless medium.

Contributions of the Paper

The main contributions of the present paper are two fold

1) *Continuous Space-Time Model*: In Section II, we define precisely the mathematical model of the network along with the assumptions we are imposing for the mathematical analysis. This model is one of the contributions of the present paper as it captures both the stochastic interactions in space and time. In Section III we state the main mathematical results of our paper. In subsection III-A we give an *exact* characterization of the time-ergodicity criterion. In section III-B, we prove the intuitive result that, when it exists, the steady-state point process in our model exhibits a form of clustering (made precise later), which is detrimental to performance as it creates higher interference powers at typical receivers than in a network with complete independence. We provide proof of ergodicity criterion in the Appendix (Section VIII) which requires the use of Point-Process theory and in particular Palm calculus and stochastic coupling arguments.
2) *Approximate Formulas for Delay and Insights*: In Section IV, we propose two formulas to approximately compute the mean-number of links per-unit space and the average delay of a link. The simplest heuristic is a first order Poisson approximation which relies on a single intensity parameter and hence cannot take clustering into account. There is numerical evidence that this

first order approximation is also a bound which overestimates the system performance. This bound is tight for light and heavy traffic, but loose in the intermediate cases. Hence we propose another heuristic, which is a second order cavity type approximation of the second moment measure [15] of the steady-state point process. This heuristic we find through simulations to work very well in all regimes. In Section V, we propose and evaluate through simulations certain other heuristics for performance to get further insight into the dynamics in our model. Finally, we conclude by remarking how one can propose and analyze similar models in the case of MIMO channels.

II. System Model - Birth-Death Model for Wireless Flows

In this section, we precisely describe the mathematical model of spatial birth and death network which we later analyze.

A. Spatial Domain

The wireless links considered in this setup are transmitter-receiver pairs. The network at any point of time consists of a certain number of transmitters each transmitting to its own unique intended receiver. This is also commonly referred to as the "dipole-model" of a D2D ad-hoc wireless network.

The wireless links live in $\mathbf{S} \subset \mathbb{R}^2 = [-Q, Q] \times [-Q, Q]$, a square region of the Euclidean plane where Q is a large but fixed *finite* constant. To avoid edge effects, we identify the opposite edges of the square and wrap it around to form a torus. We denote by $|\mathbf{S}|$ as the area of the region $\mathbf{S}$ which is $4Q^2$. We present the mathematical analysis assuming $\mathbf{S}$ is a square torus as it makes exposition of proof ideas easier.

B. Links and Traffic Arrival Process

The links arrive into the network as a stationary marked space-time process on $\mathbf{S} \times \mathbb{R}$ with intensity λ. This marked point-process on $\mathbf{S} \times \mathbb{R}$ is denoted by $\mathcal{A}$. An atom p of $\mathcal{A}$ represents the receiver and is denoted by (x_p, b_p). $x_p \in \mathbf{S}$ denotes the spatial location of receiver p and $b_p \in \mathbb{R}$ denotes the time of arrival into the network of receiver p. Hence, one can represent the point process $\mathcal{A}$ as $\mathcal{A} = \sum_{p \in \mathbb{Z}} \delta_{(x_p, b_p)}$, where $\delta_{(x,b)}$ refers to the Dirac-mass at $(x, b) \in \mathbf{S} \times \mathbb{R}$. To each point p of $\mathcal{A}$, we associate a vector mark of (y_p, L_p), where $y_p \in \mathbf{S}$ and $L_p \in \mathbb{R}^+$, where y_p refers to the location of the transmitter of receiver p and L_p denotes the file-size which the transmitter of p wants to send to the receiver of p. We refer to the pair $(x_p; y_p)$ as link p whose receiver is in location x_p and transmitter in location y_p.

The set of links present or *alive* in the network at time t is denoted by ϕ_t i.e. $\phi_t = \{(x_1; y_1), ..., (x_{N_t}; y_{N_t})\}$, where N_t is the number of links alive in the network at time t. The exact dynamics describing which links are present at a particular time t will be specified in the sequel. More formally, $\phi_t = \sum_{i=1}^{N_t} \delta_{x_t}$ is a point-process on $\mathbf{S}$ of receivers marked with the location of their transmitters. We use the terminology "*configuration of links*" to refer to a marked point-process on $\mathbf{S}$ (atoms representing the receiver locations) with its marks (representing its corresponding transmitter locations) in $\mathbf{S}$. We denote by $\phi^T = \{y_1, ..., y_{N_t}\}$, the point-process corresponding to the marks of the configuration of links ϕ (for e.g. ϕ_t^T corresponds to the point-process of transmitters present in the network at time t) and ϕ^R to denote the unmarked point-process of the configuration ϕ (for e.g. ϕ_t^R refers to the point-process of receivers present at time t).

This arrival process can be seen as an incarnation of links initiating communication in a IoT or a D2D network for instance. When it has a file to transmit (which comes randomly in time), a node "switches on" and initiates contact with its receiver. Across times, these devices may move around (especially in a D2D setting) and when the next file arrives to be transmitted, the spatial location can be different from that of the previous file arrival instant in this node. When viewed from outside,

these random communication initiations can be thought of as a point-process on space-time which we model as the link arrival process.

C. Data Rate

The transmitter of each link p has a file of size L_p measured in units of bits which needs to be communicated to its receiver. The transmitter sends this file to its receiver at a time varying rate given by the instantaneous Shannon rate. More precisely, given a configuration of links ϕ and a link $(x; y) \in \phi$, define the following function

$$R(x, \phi) = C \log_2 \left(1 + \frac{1}{N_0 + I(x, \phi)}\right). \tag{1}$$

In the above expression, C is a constant with units in bits per unit time, N_0 denotes the thermal noise power at the receiver and $I(x, \phi)$ denotes the interference seen at location x due to configuration ϕ. The interference at location x is the sum of attenuated powers from the transmitters in $\phi^T \setminus \{y\}$. The attenuation is distance dependent and is encoded by a "path-loss function" $l(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$. This then gives

$$I(x, \phi) = \sum_{u \in \phi^T \setminus \{y\}} l(||x - u||), \tag{2}$$

where $||x - y||$ refers to the Euclidean distance between the points x and y in the torus $\mathbf{S}$. Note that the interference at any receiver is the sum of attenuated powers from all other transmitters other than the transmitter of this tagged receiver under consideration. Denote by a the constant (which can possibly be infinite) $a = \int_{x \in \mathbf{S}} l(||x||)dx$.

Some common examples of path-loss functions are

- $l(r) = r^{-\alpha}$ with $\alpha > 2$ called the "power-law path-loss" model.
- $l(r) = (r + k)^{-\alpha}$ where k is a constant is commonly called the "bounded path-loss" model.

In our analysis however, we remain general and do not explicitly assume a particular form for the function $l(\cdot)$. Equation (1) is the Shannon formula for the Gaussian SISO (Single Input Single Output) channel with signal power 1 and the noise coming in from interference and thermal noise [16]. We will comment on extensions to MIMO channels in Section VI.

In Equation (1), we did not consider the effect of random channel fading. However, one can easily model the effect of fast fading by defining the rate-function as

$$R^{(f)}(x, \phi) = \mathbb{E}_h \log_2 \left(1 + \frac{h_{xy}}{N_0 + \sum_{t \in \phi^{Tx} \setminus \{y\}} h_{xt} l(||t - x||)}\right), \tag{3}$$

where h_{xy} and h_{tx} are independent random-variables representing the value of the fading power between the different transmitters and receivers and the expectation is with respect to this random vector of fades h. All of our theoretical results extend to this case but with a bit more notation and computation cost. Since the proof ideas are the same, we decided to only discuss the model without fading.

D. The Dynamics

This setup now allows one to precisely define the network dynamics. A link arriving with receiver in location $x_p \in \mathbf{S}$ and its transmitter at location $y_p \in \mathbf{S}$ at time t_p with file of size L_p leaves the network at time d_p given by the following recursive definition

$$d_p = \inf \left\{ t > b_p : \int_{u=t_p}^{t} R(x_p, \phi_u) du \geq L_p \right\}. \tag{4}$$

In the above equation, ϕ_u denotes the point process of all links "alive" at time u i.e. $\phi_u^{Rx} = \sum_{p \in \mathbb{Z}} \delta_{x_p} \mathbf{1}_{\{u \in [b_p, d_p]\}}$ and $\phi_u^{Tx} = \sum_{p \in \mathbb{Z}} \delta_{y_p} \mathbf{1}_{\{u \in [b_p, d_p]\}}$ where δ_x denotes to the Dirac-measure at

location $x \in \mathbf{S}$. We refer to the time instant b_p as the "birth" time of link p and d_p as the "death" time of link p. This is the justification for calling this dynamics a "spatial birth-death" model, i.e. transmitter-receiver pairs are "born" at time b_p and "die" at time d_p and leave the network.

This model is the wireless analog of the "flow-level" model introduced by Massoulie and Roberts [10] to evaluate and study wired networks, particularly the Internet. The flow-model in the present paper is based on a more precise modeling of the wireless interactions compared to the standard conflict graph model of interference. The spatial birth-death model can also be viewed as a "dynamic" version of the model considered in [17], namely the Gaussian Interference channel with point-to-point (ptp) codes. In our model, each link or a "flow" is a guassian ptp channel using a ptp codebook (as made evident in the rate-formulation in Equation 1). It was shown in [17], that one can consider other schemes such as Successive Interference Cancellation or Joint Optimal Decoding to get strictly better performance than considering IAN in cases of static links that use ptp codes. We however only study the dynamic version of the IAN case and leave the other cases for future work.

E. Mathematical Assumptions

All the analysis and results rely on the following assumptions on the system model presented in the previous section.

1) The link arrival process is a time-space stationary Poisson Point Process of intensity λ. The probability of an arrival of a receiver in an infinitesimal location dx in an infinitesimal time interval dt is $\lambda dxdt$.
2) The file sizes of each transmitter are i.i.d. and exponentially distributed with mean L bits.
3) The transmitter location y of a receiver at x is assumed to be located uniformly and independently of everything else on the perimeter of a ball of radius T centered at x.
4) The thermal noise power $N_0 > 0$ is a fixed constant.
5) The path-loss function is bounded and non-decreasing. This is a reasonable assumption since energy is only dissipated on traveling through space and the received energy can be no larger than the transmit energy.

These assumptions (especially the statistical ones) are imposed primarily for mathematical tractability. It is well known, at least in the context of the Internet, that file sizes are Pareto [18] and it would make modeling sense to assume heavy-tailed file sizes. We will relax the statistical assumption on exponential file-sizes in the simulation studies.

In our model, we always have unit signal strength (as reflected in Equation (1)), independent of the link length T. Hence, our mathematical problem statement is independent of T and thus, our results also do not depend on T and hold for all T. The interpretation is that the link-lengths are those obtained after a form of "zoom" on the network, and that this zoom does not affect the signal strength. Of special interest to us will be the scenario when $T = 0$. This corresponds to the scenario when one "zooms out" and the link-lengths are much smaller compared to the ambient network size (i.e. Q is very large). In this special case, the point process ϕ_t is simple and unmarked since the transmitter and receiver locations are identical, and the signal power is nevertheless still one. The interference function at a point x from configuration ϕ is then $I(x, \phi) = \sum_{y \in \phi \setminus \{x\}} l(||y - x||)$. This is of special interest as it yields a much clearer picture of the dynamics, especially that of clustering (defined later). We however observe that none of our mathematical theorems or simulation results depends on the value of T. Hence, we assume this in certain proofs to provide a clearer picture.

The statistical assumptions, namely the Poisson arrival process and i.i.d. exponential file sizes imply that the process ϕ_t is a continuous time measure-valued Markov Chain on the state space of marked simple counting measure on $\mathbf{S}$ denoted as $\mathbf{M}(\mathbf{S})$ [15]. More precisely, the process ϕ_t is a piece-wise constant jump Markov Process i.e., from a time t, the next *change* in the configuration will occur after an exponentially distributed time duration with rate $\lambda|S| + \frac{1}{L}\sum_{x \in \phi_t} R(x, \phi_t)$. This

Notation	Description
ϕ_t	Point-process of receivers alive at time t marked with their corresponding transmitter locations
$\phi_t(\mathbf{S})$	The number of links alive at time t
ϕ_t^{Tx}	The point-process of transmitter locations at time t
ϕ_t^{Rx}	The point-process of receiver locations at time t
ϕ_0	The steady state marked point-process corresponding to ϕ_t
$\phi_0(\mathbf{S})$	The random variable denoting the number of links in steady-state
$\mathbb{E}^0_{\phi_0}$	The Palm probability measure with respect to ϕ_0
$K_\phi(r)$	The Ripley K-function for point process ϕ
T	The link length in the model
β	The intensity of the point process ϕ_0
a	$\int_{x\in\mathbf{S}} l(\|\|x\|\|)dx$

TABLE I: Table of Notation

interpretation follows since births occur at the epochs of an exponential clock with rate $\lambda|S|$ and the death rate of any receiver x in configuration ϕ is $\frac{1}{L}R(x,\phi)$ which is independent of everything else. The assumption $Q<\infty$ ensures that ϕ_t is a piece-wise constant jump process. Extending the analysis of stability to the case of $\mathbf{S}=\mathbb{R}^2$ is way more challenging and is left for future work. The large torus is meant to emulate the Euclidean space. The fact that it is similar to the Euclidean space (in terms of interference field and hence birth and death dynamics) justifies our use of the Palm calculus of the Euclidean space rather than that of the torus in some derivations.

The first natural question we ask about ϕ_t is that of time ergodicity which we address in the next section. Time ergodicity implies that the process ϕ_t admits an unique steady-state in which the links form a stationary and space-ergodic point process on $\mathbf{S}$. Moreover, since $\mathbf{S}$ is a compact set, the stationary-regime when it exists will put only finitely many points in $\mathbf{S}$ at any given instant almost-surely. Denote by ϕ_0 the steady-state point-process of links i.e. the links that are "alive" or active in steady-state. ϕ_0 is a point-process on $\mathbf{S}$ with atoms representing the locations of receivers and marks representing the relative transmitter locations.

Denote by β the density of links present in the network in steady-state (assuming it exists). More formally, β denotes the intensity of the receiver point-process ϕ_0^R (which is the ground point process of ϕ_0) on $\mathbf{S}$ when the dynamics is in steady state. Note that the intensity of the transmitter point-process ϕ_0^T in steady-state is also β since every receiver in the model has exactly one transmitter. The distribution of the relative location of the transmitter of a typical receiver of ϕ_0^R is uniform on the perimeter of a ball of radius T around this receiver. However, the transmitter locations across different receivers of ϕ_0^R are not independent due to the correlation (clustering) induced by the dynamics.

The interpretation of time ergodicity is also connected to the phase-transition of mean delay. Little's law for this dynamics yields $\beta=\lambda W$, where W is the average sojourn time of a typical link i.e., $W=\mathbb{E}[d_0-b_0]$; which follows from PASTA [19]. The process ϕ_t being time ergodic is equivalent to asserting that $W<\infty$, i.e. finite mean delay for a typical link in the network.

III. Main Theoretical Results

The main theoretical results of our paper are on the time-ergodicity (or stability) conditions of the dynamics ϕ_t and on a certain structural characterization of the steady-state point process of ϕ_t whenever it exists. As noted earlier, this result holds for all T (since the mathematical problem statement is independent of T) and in particular for the case of $T=0$ as well when there is an unmarked point process ϕ_t. The proofs of the theorems are presented in the Appendix. Recall the notation that $a=\int_{x\in\mathbf{S}} l(||x||)dx$.

A. Stability Criterion

We state our main theoretical results on the stability criterion (i.e. time ergodicity) of the dynamics.

Theorem 1. *If $\lambda > \frac{C}{\ln(2)La}$, then the Markov Chain ϕ_t admits no stationary regime.*

We see from the proof (in Section VIII) that this theorem only needs the weaker assumption that $l(\cdot)$ be such that $l(r) < \infty$ for all $r > 0$. This indeed is a weaker assumption than assuming that the function $l(\cdot)$ is bounded. Thus, we have as immediate corollary to this theorem:

Corollary 1. *For the path-loss model $l(r) = r^{-\alpha}$, $\alpha > 2$, for all $\lambda > 0$, and all mean file sizes, the process ϕ_t admits no stationary-regime.*

Proof. This follows since the integral $\int_{x\in\mathbf{S}} l(||x||)dx$ diverges for the function $l(r) = r^{-\alpha}$ for all $\alpha > 2$. □

The next result provides the sufficient condition for time ergodicity i.e. existence and uniqueness of a stationary regime.

Theorem 2. *If $\lambda < \frac{C}{\ln(2)La}$, then the Markov Chain ϕ_t is time ergodic, i.e. has an unique stationary regime.*

We note that the above theorems are valid even in the case of fading if one used the rate-function in Equation (3) with the fades being unit-mean i.i.d. random variables. The two theorems identify the exact critical arrival rate λ for ergodicity as $\lambda_c = \frac{C}{L\ln(2)a}$. We however refrain from studying the critical case as it is technically more subtle.

In the sequel, whenever we refer to ϕ_0, we implicitly assume ϕ_t is ergodic, i.e. the condition $\lambda < \frac{C}{\ln(2)La}$ holds.

B. Clustering

In this section, we state the main structural characterization of the steady-state point process ϕ_0 when it exists i.e. when $\lambda < \frac{C}{L\ln(2)a}$. We need the following definition of clustering.

Definition 1. *(CLUSTERING) Let ϕ be a stationary configuration of links, i.e. it is a stationary marked point-process on $\mathbf{S}$ with its marks in $\mathbf{S}$. Then ϕ is said to be clustered if for all bounded, positive, non-decreasing functions $f(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$, the following inequality holds*

$$\mathbb{E}^0_\phi[F(0,\phi)] \geq \mathbb{E}[F(0,\phi)], \tag{5}$$

where F is the shot-noise defined as follows. For any atom (receiver) $x \in \phi$ with its corresponding mark (transmitter) $y \in \mathbf{S}$, the shot noise $F(x,\phi) := \sum_{T\in\phi^{Tx}\setminus\{y\}} f(||T-x||)$.

Theorem 3. *If the dynamics ϕ_t is ergodic, then the steady state point process ϕ_0 is clustered.*

By substituting $f(\cdot) = l(\cdot)$ in Equation (5) , we get that the mean of the interference measured at any uniformly randomly chosen receiver in the steady-state point process (this is the interpretation of the Palm probability) is larger than the mean of the interference measured at any uniformly randomly chosen location of space in $\mathbf{S}$.

To get an understanding as to why the above definition is a form of clustering, consider the case $T = 0$ which gives a clearer picture. In this case, Theorem (3) gives a clustering comparison of ϕ_0 with a Poisson Point Process (PPP) of same intensity. Let ψ be a PPP of the same intensity as ϕ_0. Then, from Slivnyak's theorem (Theorem 1.4.5, [3]), one can rewrite the inequality in (5) as

$$\mathbb{E}^0_{\phi_0}[F(0,\phi_0)] \geq \mathbb{E}^0_\psi[F(0,\phi_0)], \tag{6}$$

where $\mathbb{E}^0_\psi[F(0,\phi_0)] = \mathbb{E}[F(0,\psi)]$ follows from Slivnyak's theorem which is equal to $\beta\int_{x\in\mathbf{S}} f(||x||)dx$ from Campbell's Theorem (Theorem 1.4.3, [3]). Slivnyak's theorem essentially gives that the PPP has no clustering i.e. the Inequality 5 is an equality. Hence, we automatically have a shot noise

comparison of the steady state point process ϕ_0 with a PPP.

The comparison with a PPP also gives us a comparison of Ripley K-function [20] of ϕ_0 with that of a PPP. The Ripley K-function $K_\phi(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ of a point-process ϕ is defined as $K_\phi(r) = \frac{1}{\beta}\mathbb{E}^0_\phi[\phi(B(0,r)) - 1]$ where β is the intensity of ϕ and $\mathbb{E}^0_\phi$ is the Palm probability measure of ϕ. This function can be interpretted as the mean number of points (scaled by the intensity of the point-process) within distance r to the origin *conditioned* on a point of ϕ to be present at the origin. Ripley K-function is commonly used in statistical analysis of point-patterns to identify if an empirical data-set exhibits statistical clustering [20]. Based on the shot-noise comparison with a PPP, we have the following corollary.

Corollary 2. *Assume ϕ_t is in steady-state and $T = 0$. Denote by β to be the intensity of ϕ_0 and ψ to be a PPP on $\mathbf{S}$ with intensity β. Then, $K_{\phi_0}(r) \geq K_\psi(r)$.*

Proof. Consider $f(x) = \mathbf{1}(x \leq r)$ in Theorem 3. □

We will use Ripley K-function in the simulations to compare the point process ϕ_0 with a PPP to derive a bound on the intensity β of ϕ_0 as a function of λ, L and $l(\cdot)$.

Intuitively, it is not surprising to expect a clustered point-process in steady state. An arriving link gets lower rate if it is in a crowded area of transmitters, due to interference. This arriving link also causes more interference to the cluster of links already present thereby causing more interference and slowing everyone down. This reinforcement of service slowdown is actually the fundamental reason making the system always unstable in the power law attenuation function case. More generally, when ϕ_t is sampled in steady-state, it is expected to be clustered which Theorem 3 formalizes. A snapshot of the point-process ϕ_0 is presented in Figure 4 which gives a visual understanding of the clustering.

IV. Performance Analysis - Steady State Formulas

In this section, we propose two heuristic formulas for β the intensity of the point process ϕ_0 as a function of λ. Note that a heuristic formula for β gives a heuristic formula for mean delay W through Little's Law ($\beta = \lambda W$).

We propose two formulas - β_f called the *Poisson Heuristic* and β_s called *Second-Order heuristic* to approximate β the intensity of the steady-state point process ϕ_0. We show that subject to a natural conjecture (Conjecture 1), β_f is a lower bound on β. We see from simulations however that β_s is a much better approximation of β compared to β_f. Both these formulas are derived based on approximately evaluating the following Equation we established in (19)

$$\lambda L = \beta \mathbb{E}^0_{\phi_0}\left[\log_2\left(1 + \frac{1}{N_0 + I(0)}\right)\right]. \tag{7}$$

The Poisson Heuristic

The Poisson heuristic formula β_f is given by the largest solution to the following fixed point equation

$$\lambda L = \frac{\beta_f}{\ln(2)} \int_{z=0}^{\infty} \frac{e^{-N_0 z}(1 - e^{-z})}{z} e^{-\beta_f \int_{x \in \mathbf{S}} (1 - e^{-z l(||x||)}) dx} dz. \tag{8}$$

This formula is obtained by approximating the expectation in Equation (7) by assuming the following "*Independent Poisson heuristic*". We assume that ϕ_0 is an independently marked Poisson-Point process with the transmitter locations of different receivers in ϕ_0 being independent. Since the transmitter locations are assumed to be independent, the process ϕ_0^{Tx} will also be a PPP in this Poisson heuristic. We state the following lemma without proof from [21] which is useful in computing the expectation under the Poisson assumption.

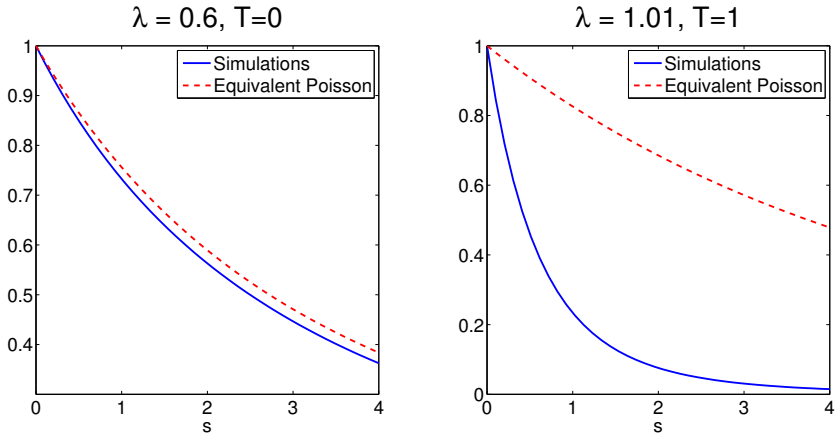


Fig. 1: A plot comparing the functions $\mathbb{E}^0_{\phi_0}[e^{-sI(0;\phi_0)}]$ and $\mathbb{E}^0_{\psi}[e^{-sI(0;\psi)}]$, for $l(r) = (r+1)^{-4}$.

Lemma 1. *Let X, Y be non-negative and independent Random Variables. Then,*

$$\mathbb{E}\left[\ln\left(1+\frac{X}{Y+a}\right)\right] = \int_{z=0}^{\infty} \frac{e^{-az}}{z}(1-\mathbb{E}[e^{-zX}])\mathbb{E}[e^{-zY}]dz.$$

We can then explicitly compute the expectation in Equation (7) as follows

$$\begin{aligned}\lambda L &= \beta_f \mathbb{E}^0_{\psi}\left[\log_2\left(1+\frac{1}{N_0+I(0)}\right)\right] \\ &\overset{(a)}{=} \beta_f \mathbb{E}_{\psi}\left[\log_2\left(1+\frac{1}{N_0+I(0)}\right)\right] \\ &\overset{(b)}{=} \frac{\beta_f}{\ln(2)}\int_{z=0}^{\infty}\frac{e^{-N_0 z}(1-e^{-z})}{z} e^{-\beta_f \int_{x\in\mathbf{S}}(1-e^{-zl(||x||)})dx}dz,\end{aligned}$$

where ψ is a Poisson Point Process on $\mathbf{S}$ with intensity β_f. The equality (a) follows from Slivnyak's theorem and the equality (b) follows from Lemma 1 and the formula for the Laplace functional of a Poisson Point Process. The subscript f refers to the computation of the density under this Poisson heuristic. This establishes the formula in Equation (8).

We now make the following conjecture on the higher-order moment measures of ϕ_0, which we will leverage to show that β_f is a lower bound on β.

Conjecture 1. *Let ϕ_0 be the point process on $\mathbf{S}$ corresponding to the stationary distribution of ϕ_t with intensity β. Denote by ψ to be an independently marked Poisson Point Process on $\mathbf{S}$ with intensity β. The mark of any atom x of ψ is a point y drawn uniformly on the perimeter of a circle of radius T around x. Then, for any $s > 0$, we have $\mathbb{E}^0_{\phi_0}[e^{-sI(0;\phi_0)}] \leq \mathbb{E}^0_{\psi}[e^{-sI(0;\psi)}]$.*

Note that from Slivnyak's theorem we also have $\mathbb{E}^0_{\psi}[e^{-sI(0;\psi)}] = \mathbb{E}_{\psi}[e^{-sI(0;\psi)}]$. This conjecture which is validated through simulations in Figure 1, is a stronger statement on the structural characterization of ϕ_0 than stated in Theorem 3. This conjecture gives that not only the mean, but each moment of the interference measured at a typical receiver of ϕ_0 is larger than that measured at a typical receiver of an equivalent PPP. We saw in Corollary 2 that the Ripley K-function for ϕ_0 is larger than that of an equivalent Poisson Point Process of the same intensity. This says that *conditioned* on a point of ϕ_0 being present at the origin, one would expect more points to be closer to the origin in ϕ_0 than in an equivalent Poisson Point Process of the same density. The intuition for the conjecture then follows from the fact that the path-loss function $l(\cdot)$ is non-decreasing.

Proposition 4. *Subject to Conjecture (1), we have that $\beta \geq \beta_f$, where β_f is the largest solution of Equation (8).*

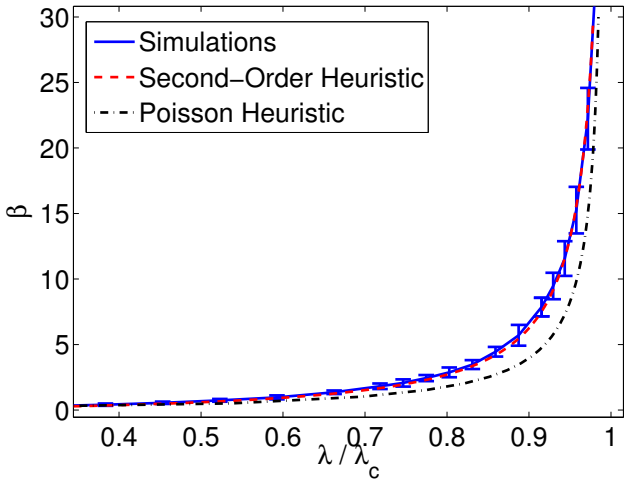


Fig. 2: The performance plot with 95% confidence interval when $T = 0$ and $l(r) = (r+1)^{-4}$.

Proof. Let $g(\beta) = \beta \mathbb{E}^0_{\phi_0}[R(0;\phi_0)]$ (where ϕ_0 has intensity β) and let $p(\beta) = \beta \mathbb{E}^0_{\psi}[R(0;\psi)]$ where ψ is a PPP on $\mathbf{S}$ with intensity β. Rate-conservation equation (7) gives that $\lambda L = g(\beta)$ and our heuristic computation is $\lambda L = p(\beta_f)$. From our conjecture and Lemma 1, we have the inequality $g(\beta) \leq p(\beta)$. The function $g(\beta) = \beta \mathbb{E}^0_{\phi_0}[R(0;\phi_0)]$ is monotone non-decreasing in β as it describes the true dynamics through the equation $\lambda L = g(\beta)$. The monotonicity of $g(\cdot)$ along with the inequality $g(\beta) \leq p(\beta)$ gives the performance bound $\beta \geq \beta_f$. □

An interesting observation here is that β_f does not depend on the link-distance T and is a bound on β for all values of T. Proposition 4 then gives $\beta_f|\mathbf{S}|$ as a lower bound on the mean number of links present in the network in steady state and $\frac{\beta_f}{\lambda}$, as a lower bound on mean-delay of a typical link.

The Poisson heuristic completely ignores the spatial clustering we established in Theorem 3 and assumes complete-spatial randomness. Since it does not account for the clustering it underestimates the typical interference seen at a receiver and therefore predicts a lower density of links. We see through simulations, that this heuristic is poor (i.e. the gap between β and β_f is large) in certain traffic regimes (Figure 2). This is not surprising as one cannot neglect the effect of spatial correlations except in asymptotic regimes of heavy and light-traffic (detailed later). Motivated by the poor performance of the Poisson heuristic in certain regimes, we propose a "second-order heuristic" β_s which takes into account the spatial correlations by considering an approximation of the second-order moment measure of ϕ_0. We see through simulations (Figure 2) that β_s is a much better approximation of β than β_f in all traffic regimes.

Second-Order Heuristic

We propose a heuristic formula β_s for approximating β in Equation (9). For all values of T, β_s is given by

$$\beta_s = \frac{\lambda L}{C \log_2\left(1 + \frac{1}{N_0 + I_s}\right)}, \tag{9}$$

where I_s is the smallest solution of the fixed-point equation

$$I_s = \lambda L \int_{x \in \mathbf{S}} \frac{l(||x||)}{C \log_2\left(1 + \frac{1}{N_0 + I_s + l(||x||)}\right)} dx. \tag{10}$$

We call the heuristic in Equation (9) a *second-order heuristic* since it follows from an approximation of the second-order moment measure of ϕ_0 as follows. Let I_s denote the interference of a typical point at ϕ_0 and *assume* it is non-random and equal to its mean. Then, Equation (9) follows from RCL

in Equation (7). To compute I_s, we use the following approximation of the second order moment measure $\rho^{(2)}(x,y)$ of ϕ_0 as

$$\rho^{(2)}(x,y) \approx \frac{\beta\lambda L}{C\log_2\left(1+\frac{1}{N_0+I_s+l(||x-y||)}\right)}. \tag{11}$$

Intuitively, the approximation is a form of cavity approximation which can be understood as follows. Two points at locations x and y will each "see" an interference of I_s which is the interference of a typical point plus the additional interference caused by the presence of the other point. Using the above interpretation of interference, Equation (11) is a form of RCL on the pair of points at x and y. The average increase of the pair happens at rate $2\lambda\beta$ and the average decrease of the pair happens at the rate equal to the sum of rates (since file-sizes are i.i.d. exponential) received by points x and y which is approximately $2(C/L)\log_2\left(1+\frac{1}{N_0+I_s+l(||x-y||)}\right)$ from the cavity approximation. Now, using the fact that $\mathbb{E}^0_{\phi_0}[I_0] := I_s = \frac{1}{\beta}\int_{x\in\mathbf{S}} l(||x||)\rho^{(2)}(x,0)dx$, we get Equation (10) from Equation (11).

The heuristic β_s is compared against the true β and the Poisson heuristic β_f in Figure 2. The second-order heuristic performs much better compared to the Poisson-heuristic as it takes into account some notion of spatial correlations which the Poisson heuristic completely ignores.

V. Simulation Studies

We perform simulations to gain a finer understanding of our model which can pave way for future mathematical research. We see that the bound in Proposition 4 is tight in the two asymptotic regimes of light and heavy-traffic where the effect of spatial correlations vanishes. We also argue that, in these two asymptotic regimes, the heuristic β_s is "close" to β_f thereby implying that β_s is also a good approximation to β. As noticed in Figure 2, β_f is a poor approximation to β compared to β_s in the intermediate traffic-regime which we further highlight in this section.

To explore the impact of spatial correlations, we study the tails of delay of a typical link and correlation between delays of different links in space. We observe that our model exhibits marked difference in terms of tail delay behavior from that of an equivalent queuing system which is obtained by a "spatial-fluid" approximation. We conclude from these studies that although our model resembles that of a queue (for e.g. the dynamics satisfies Little's Law), there are significant differences due to the spatial correlations, which in hindsight is not so surprising. We finally perform simulations with heavy-tailed file size distribution and observe qualitatively the same phenomena as seen under exponential file-size distribution. We state our simulation results as claims which are not formal conjectures, but are meant to provide a starting point for future research.

A. Simulation Setup

The path-loss function we consider is $l(r) = (r+1)^{-4}$. Although all of the results qualitatively hold for any bounded-non-increasing function, we choose this power law function due to its wide-spread popularity in modeling wireless propagation. We assume unit link-length $T = 1$ unless otherwise mentioned. We however note that all the qualitative results carry over for any value of T including the case of $T = 0$. The pictures of point-process and the Ripley K-functions we test are those corresponding to the receivers.

B. Tightness of $\beta_{\mathbf{f}}$

We study the bound in Proposition 4, by empirically noticing how much K_{ϕ_0}, the Ripley K-function of ϕ_0 deviates from that of an equivalent PPP denoted by K_{PPP}. The two Ripley K-functions being almost identical implies that the steady-state is "almost" Poisson and thereby the bound in Proposition 4 is good. On the other hand, if there is significant deviation between the two Ripley K-function,

then the bound is poor. We know from Corollary 2 that $K_{\phi_0}(r) \geq K_{\mathrm{PPP}}(r)$ for all $r \geq 0$. Here, we are interested in seeing how large this difference can be.

To plot the Ripley K-functions, we simulated the Markov chain ϕ_t in forward time for a long time to obtain a single sample of the steady-state ϕ_0. We used the Spatstat package in R [22] to perform spatial statistics and plot K_{ϕ_0}. A single sample is sufficient as we take a large enough space (i.e. large $\mathbf{S}$) so that a single sample of ϕ_0 has about 500 points. Due to spatial ergodicity of ϕ_0, we get a smooth estimate of the K-function from a single sample.

We observe in Figures 3a, 3b and 3c, that the functions K_{ϕ_0} and K_{PPP} are very close in heavy and light traffic and are very different in intermediate traffic. The Heavy traffic corresponds to the scenario when λ is very close to the critical λ_c and the light traffic corresponds to the case when λ is very "close" to 0. We do not rigorously demarcate the exact space-time scaling needed to define the two asymptotic limiting regimes as it is beyond the scope of this paper.

Claim 1. *ϕ_0 is almost Poisson in light-traffic. Moreover, the delay of a typical link converges weakly to an exponential distribution with mean $L\log_2\left(1+\frac{1}{N_0}\right)^{-1}$ as $\lambda \to 0$.*

In the light-traffic regime, λ is very "small" compared to L, and thereby β is also "small". This then implies that the distribution for the interference $I(0,\phi_0)$ is close to 0 thereby making the interaction between the points almost negligible. The rate-function can then be approximated as $R(x,\phi) \approx \log_2\left(1+\frac{1}{N_0}\right)$ and the dynamics resembles that of a spatial $M/M/\infty$ queue whose stationary spatial distribution is a PPP. The intensity β in this regime is $\beta_l = \lambda\left(\log_2\left(1+\frac{1}{N_0}\right)\right)^{-1}$. The subscript l refers to the density computation in the interaction-less approximation. Figure 3c shows evidence that ϕ_0 exhibits very little clustering in this regime and is "close" to a PPP.

Claim 2. *In the heavy-traffic regime, ϕ_0 is almost Poisson, i.e. the effect of clustering vanishes as $\lambda \to \lambda_c$.*

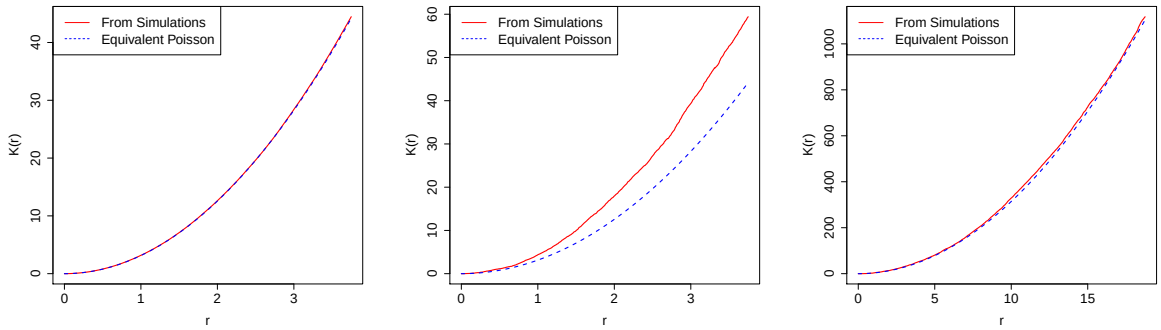
The intuition behind the heavy-traffic behavior is that as λ approaches λ_c, the stationary distribution is very dense i.e. β is large. Hence, the interference of a typical arriving link is mainly dominated by the local geometry which does not change much during the life-time of the typical link. This indicates that the dynamics then behaves very similarly to a heavily loaded $M/M/1$ Processor Sharing (PS) queue and the correlation across space is negligible in this regime. Moreover, it is easy to see that $\lambda = \lambda_c$ is an asymptote for Equation (8) i.e. as $\lambda \to \lambda_c$, $\beta_f \to \infty$. This further strengthens the belief that the stationary distribution is close to Poisson in the heavy-traffic regime as it predicts the correct stability boundary. Making this claim rigorous or even just state a mathematical conjecture is quite challenging and would require an appropriate scaling of space and time similar to the diffusion scaling considered for a single server PS queue [23].

C. Tightness of $\beta_{\mathbf{s}}$

We argue here that in both the low and heavy-traffic regimes, the approximation β_s is close to β_f and is hence a good approximation of the true β. In low-traffic, as $\lambda \to 0$, the smallest solution of Equation (10) tends to 0 and hence the formula for $\beta_s \approx \frac{\lambda L}{C\log_2\left(1+\frac{1}{N_0}\right)}$. This from Claim 1 gives that β_s and β_f predict the same value in low-traffic. In high-traffic regime, as $\lambda \to \lambda_c$, the value if I_s from Equation (10) is very high. Thus, the second-order moment-measure approximation $\rho^{(2)}(x,y)$ in Equation (11) is almost constant i.e. does not depend of the actual values of x and y as $l(\cdot)$ is a bounded function. This implies that the effect of clustering vanishes in this heuristic and hence is close to β_f.

D. Intermediate Clustered Regime

In the intermediate regime, the Poisson approximation is poor and the steady state-point process is quite clustered (see Figures 4 and 3b) i.e. K_{ϕ_0} is much larger than K_{PPP}. However, we see from



(a) $\lambda = 1.40$ (b) $\lambda = 0.99$ (c) $\lambda = 0.2$

Fig. 3: Plot comparing the Empirical Ripley K-function K_{ϕ_0} with that of an equivalent PPP. The path loss function is $l(r) = (r+1)^{-4}$, $T = 1$. The critical $\lambda_c = 1.42$. This shows that there is little clustering in the heavy and light traffic regimes but significant clustering in the intermediate regime.

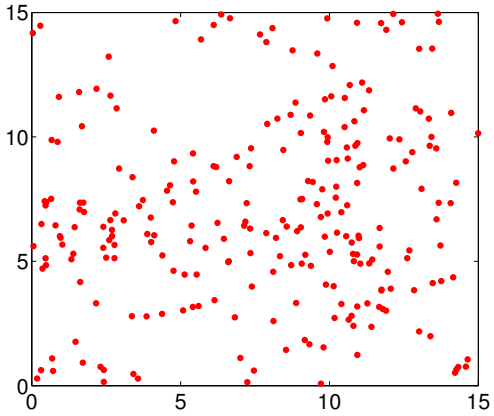


Fig. 4: A sample of ϕ_0 when $\lambda = 0.99$ and $l(r) = (r+1)^{-4}$. This is a visual representation of the clustering of points.

Figure 2 that the second-order heuristic β_s performs much better than the Poisson heuristic in this regime as it takes into account some form spatial-correlations at-least upto second-order moment measure of ϕ_0. However, Figure 4 which shows a snapshot of ϕ_0 which is a clustered process is very interesting as it indicates interesting properties of higher order moment measures. One observes for instance "*filaments*" of points which are locally directional in-spite of the fact that the dynamics is isotropic. Such behavior indicates that the higher order moment measures of ϕ_0 (of order greater than 2) may have interesting properties which we cannot capture neither in Theorem 3 nor in the heuristic β_s. Understanding the higher order moment measure of ϕ_0 can also aid in proposing a *provably* better performance bound in this intermediate regime. Studying these higher order moment measures of ϕ_0 will be a very interesting and challenging direction of research.

E. Delay Tails

To get a heuristic understanding of the delay tails, one would be tempted at first glance to approximate our model by an equivalent $M/M/1$ PS queue using a *spatial-fluid* approximation that neglects randomness in space. We see through simulations, that any approximation that neglects spatial interactions will predict much larger delays for a typical link than the true delays in our model.

Claim 3. *The delay tails in our model are exponential and have a faster decay than that of an equivalent $M/M/1$ PS queue obtained by a "spatial-fluid" approximation.*

An equivalent $M/M/1$ PS queue approximation has the following parameters - arrival rate λ, service requirement of mean L and service capacity of the server λ_c which is split equally among all customers in the queue. Such a PS queuing model is equivalent to a first-order approximation where the spatial randomness vanishes and a point in steady-state receives rate of $C\log_2\left(1+\frac{1}{N_0+\beta a}\right)$ where β is the density of points in the steady-state. Hence, the quantity $\frac{C}{\ln(2)a}$ (which is an upper bound on the total rate given to all points i.e. $C\beta\log_2\left(1+\frac{1}{N_0+\beta a}\right)\leq\frac{C}{\ln(2)a}$) can be seen as the maximum service capacity of the spectrum in $\mathbf{S}$ which is equally shared by all links accessing the spectrum. Another simple picture as to why the above $M/M/1$ PS queue is a simple heuristic is to observe that this queue corresponds to the scenario when one ignores spatial interactions among the arriving points and assumes that the total spectrum "capacity" of λ_c is shared equally among all the links sharing the spectrum in $\mathbf{S}$. Hence, the mean-delay under the $M/M/1$ - PS model for a typical point is $\frac{L}{\lambda_c-\lambda}$ and the stability criteria for this queue is the same as that for our spatial model. However, we note from simulations (Figure 6) that the delay tails predicted by the heuristic $M/M/1$ queue which completely ignores spatial interactions are much larger than those observed in our model.

The poor performance of the queuing approximation can be understood by studying the correlation between the delays of different links. In Figure 7, we plot the correlation between the delay experienced by two links arriving at the *same time* as a function of their distance. We consider the $T=0$ case and hence the distance between two links is just the distance between the two points. Numerically, we plotted Figure 7 by first sampling a steady-state point process (by running the Markov Chain ϕ_t for a long time) and then introducing two additional links to this sample with independent file-sizes. We then run the dynamics from this state until the two additional links die and then compute the correlation between their delays

We see from Figure 7 that as the distance between the two links increases, the delays of the two links are almost uncorrelated even though they arrive at the same time. This indicates that, two links arriving at the same time will be almost oblivious to each other and will each roughly receive independent service if they arrive far enough apart in space. This is unlike in the $M/M/1$ - PS queue approximation where two customers arriving at the same time have positively associated delays as both of them will be competing for the same spectrum resource. This suggests that the spatial heterogeneity is key in extracting more "service" from the spectrum than predicted by a model which considers spectrum as a fixed quantity of good to be divided among contending links.

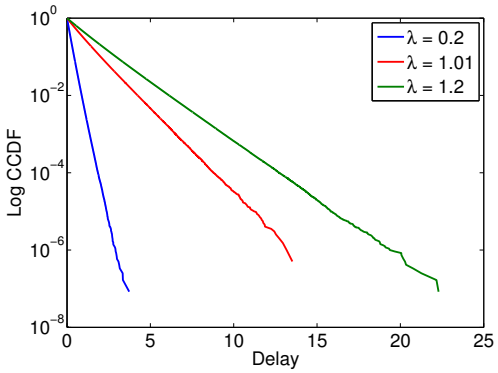


Fig. 5: Plot of logarithm of CCDF of delay.

F. Heavy Tailed File Sizes

Claim 4. *ϕ_t, with file-sizes being Pareto distributed of mean L and finite variance, admits a stationary regime with the critical λ being smaller than or equal to $\frac{C}{\ln(2)La}$.*

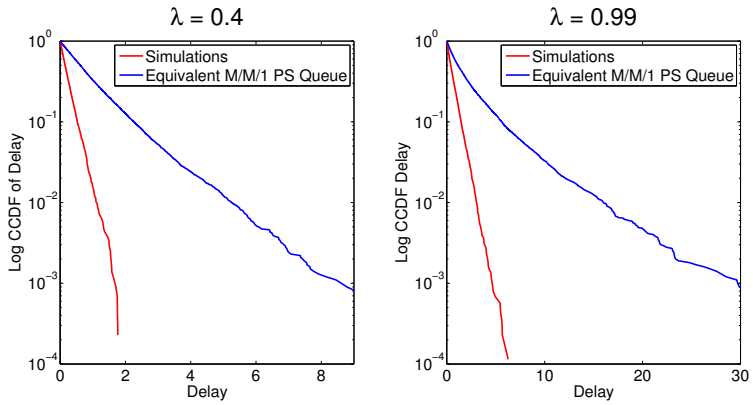


Fig. 6: Comparison of the delays with that of an equivalent M/M/1 - PS queue. The critical $\lambda_c = 1.42$.

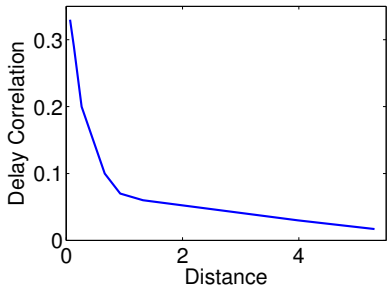


Fig. 7: Decay of delay correlation of two points born at the same time, as a function of their distance. $\lambda = 0.8$.

This model also exhibits the interesting phenomenon of prominent clustering in the intermediate traffic regime and very little to no-clustering in the asymptotic regimes of high and low traffic. Note that the term "high-traffic" in this context is somewhat loose since we do not even know exactly the stability region. With regards to delays, our model predicts tails that are stochastically dominated by the delay of a typical customer of an equivalent $M/GI/1$ PS queue (see Figure 8). The equivalent queue we simulated against had a capacity of λ_c which from Claim 4 is an upper bound on the capacity. Nonetheless, the delay predicted in our model is stochastically smaller. This observation again highlights the importance of taking into account the spatial heterogeneity in modeling the "service" provided by the spectrum.

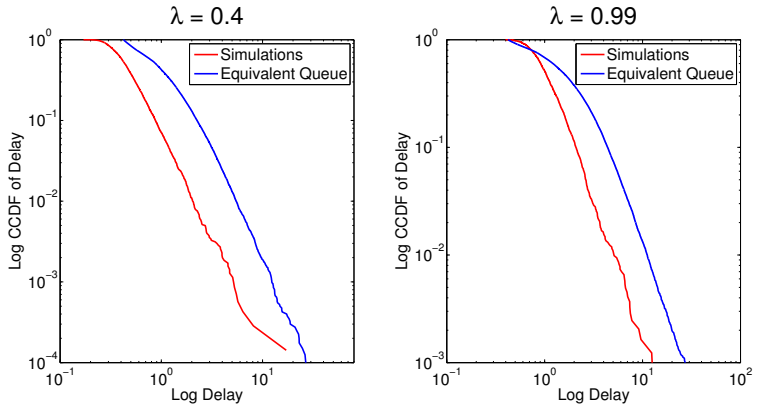


Fig. 8: Comparision of the delay under Pareto file size distribution with mean L and shape $\alpha = 2.5$.

VI. Extension to the MIMO Channel

In the previous sections, we considered the case when both the transmitter and the receiver of a link have a single antenna. In this section, we briefly highlight, an extension of the Spatial Birth-Death (SBD) model to account for the scenario when both the transmitters and receivers have multiple antennas while still treating Interference from other transmitters as Noise.

A. Generalized MIMO Framework for treating Interference as Noise

The MIMO setting is similar to the single antenna dynamics described in Section II except, the rate function in Equation (1) will be modified suitably to account for the presence of multiple antennas. In this section, we provide an extension of the rate function with fading in Equation (3) to the MIMO setting. In particular, we extend Telatar's formula [24] for MIMO channel capacity to the network case with interference treated as noise.

The Telatar's formula for capacity of a MIMO point-to-point channel with X_t transmit antennas and X_r receive antennas is given by the formula

$$\mathcal{C} = \mathbb{E}\left[\log_2\left(\det(\mathbf{I}_{X_r} + H\Sigma_S H^\dagger \Sigma_N^{-1})\right)\right], \tag{12}$$

where the expectation is with respect to the channel matrix H which is a random variable and possibly on the matrices Σ_S and Σ_N, which could in general be functions of H. The notation $H^\dagger$ is used to denote the complex conjugate of H. The matrix Σ_S is a $X_t \times X_t$ matrix denoting the outer product of the signaling vector and Σ_N is a $X_r \times X_r$ matrix denoting the outer product of the noise vector at the receiver. The matrix I_{X_r} is a $X_r \times X_r$ identity matrix. The channel capacity formula in Equation (12) captures many different scenarios such as presence or absence of Channel State Information at Transmitter (CSIT) by suitably optimizing over the correlation matrix Σ_S. For instance, in the absence of CSIT, the optimal Σ_S is deterministic, while in the presence of CSIT, the optimal power allocation is by a water-filling on the singular values of the channel matrix H and the signaling vector is along the principal components of the matrix H (Chapter 8, [25]). Thus, in the CSIT case, Σ_S will be a function of H and hence the expectation in Equation (12) will also bear on Σ_S in this case.

We present an extension of the formula in Equation (12) to a network setting by suitably defining the rate function $R(x;\phi)$ of Equation (3). Recall that the arrival process $\mathcal{A}$ is a marked PPP on $\mathbf{S}\times\mathbb{R}$ with the atoms denoting the receivers and the marks which are $\mathbf{S}$ valued denoting the location of the transmitters. We call ϕ a configuration of links if its atoms are locations of receivers and the marks of the atoms denote the corresponding transmitter locations. The notation ϕ^T is used to denote the set of transmitters or marks of the atoms of ϕ.

Proposition 5. *The generalized MIMO rate function $R(x;\phi)$ where ϕ is a configuration of links on $\mathbf{S}$ and $x \in \phi$ is an atom of the point-process ϕ is given by*

$$R(x;\phi) = C\mathbb{E}\left[\log_2\left(\det(\mathbf{I}_{X_r} + H_{xx}S_xS_x^\dagger H_{xx}^\dagger \Sigma_N^{-1}(\phi;\{H_{yx}\}_{y\in\phi\setminus\{x\}},\{S_y\}_{y\in\phi\setminus\{x\}}))\right)\right], \tag{13}$$

where the expectation is with respect to the i.i.d. collection of channel random matrices $\{H_{xy}\}_{y\in\phi}$ and the i.i.d. collection of signaling vectors $\{S_y\}_{y\in\phi}$. $\mathbf{I}_{X_r}$ denotes a $X_r \times X_r$ identity matrix. The $X_t \times 1$ vector S_x, denotes the signaling vector of the transmitter located at u whose corresponding receiver is at location x. The vector S_x may or may not depend on H_{xx} depending on whether CSIT is present or absent, but nonetheless are i.i.d. across x. The matrix Σ_N which is a $X_r \times X_r$ matrix is the outer product of the interference vector plus noise, i.e. $\Sigma_N = N_0\mathbf{I}_{X_r} + (\sum_{y\in(\phi^T\setminus\{u\})}\sqrt{l(||x-y||)}H_{yx}S_y)(\sum_{y\in(\phi^T\setminus\{u\})}\sqrt{l(||x-y||)}H_{yx}S_y)^\dagger$.

Proof. For a static and deterministic configuration of links ϕ, and for a receiver at location $x \in \phi$ with its corresponding transmitter at location u (referred to as the link in consideration in this proof), we need to argue that Equation (13) is the capacity of this link under fast-fading when treating interference as noise. Assume that the channel between any transmitter whose receiver is at location $a \in \phi$ and any receiver $b \in \phi$ is given by H_{ab} and is i.i.d. across a and b and equal in distribution to

a random matrix H. Denote by S_x as the $X_t \times 1$ random signaling vector of the transmitter of the link in consideration. Note that S_x could possibly depend on the channel realization H_{xx} depending on whether there is CSIT or not. The interference signal at the receiver in location x is

$$\mathcal{I}(x;\phi) = \sum_{y \in (\phi \setminus \{x\})^{T_x}} \sqrt{l(||x-y||)} H_{yx} S_y \tag{14}$$

where H_{yx} and S_y are i.i.d. and independent of each other with H_{yx} equal in distribution to H and S_y equal in distribution to S_x. Thus, the matrix Σ_N is the sum of the outer product of $\mathcal{I}(x;\phi)$ and the thermal noise co-variance matrix $N_0 \mathbf{I}_{X_r}$. Recall that the path-loss function $l(\cdot)$ denotes the attenuation in the signal power and hence, the signal itself is attenuated by $\sqrt{l(\cdot)}$. Now using the Telatar's formula of Equation (12) for the case when the noise signal is the sum of thermal noise and interference, we get Equation (13) which is the network version. □

The birth-death dynamics in the MIMO setting is then described through Equation (4) with the rate-function as given in Equation (13). The formulation in Equation (13) is the network version of channel capacity of MIMO under presence of fast fading and treating interference as noise. The above generalized framework can capture both the settings of presence and absence of CSIT by suitably defining the signaling vector S_x to be either a function of the realization of H_{xx} or a function of the distribution of H_{xx} respectively. This generalized setup of the MIMO channel also allows us to study the MISO and SIMO cases by setting X_r or X_t to 1 respectively. We however, only state the spatial birth-death dynamics with multiple antennas problem in the present paper and leave the complete characterization of the dynamics to future work. In the sequel, we discuss an example of the MIMO framework which can be analyzed as a corollary of the single antenna system.

B. Example MIMO Setting - Independent Channel Case with no CSIT

We show in this sub-section, that in the special case when the signal S_x is Gaussian with outer product equal to $\frac{1}{X_t}\mathbf{I}$ and the channel matrices are equal in distribution to a random matrix H where each entry is i.i.d. complex normal with 0 mean and unit variance, then the MIMO dynamics can be reduced to an equivalent single-antenna system and the critical arrival rate for this model can then be computed. Note that this transmission strategy where $S_x S_x^{\dagger} = \frac{1}{X_t}\mathbf{I}$ is optimal in the case when there is no CSIT, total transmit power constraint of 1 and the channel matrix being composed of i.i.d. entries (Chapter 8, [25]).

The following statistical assumptions model the independent channel MIMO system for which the critical arrival density λ_c can be computed as a corollary of the single antenna analysis.

- All channel realizations between any transmit antenna and receive antenna are i.i.d. complex normal with 0 mean and unit variance.
- For any coordinate $i \in [1, X_t]$, and any link x and at any time t, we have $\mathbb{E}[(S_x^t)_i (S_x^t)_i^*] = 1/X_t$ where the vector S_x^t is the transmitted signal by transmitter whose receiver is at location x at time t. This indicates that the total power 1 is split equally on each antenna.
- For any coordinates $i, j \in [1, X_r]$ and any two links $x \neq y$ and any time t, we have $\mathbb{E}[(S_x^t)_i (S_y^t)_j^*] = 0$. This assumption gives that the signal across antennas are uncorrelated.

Under the foregoing assumptions, the rate function in Equation (13) can be simplified to -

$$R(x,\phi) = C\mathbb{E}_h \left[\sum_{i=1}^{X_r} \log_2 \left(1 + \frac{1}{X_t(N_0 + I(x;\phi))} \sigma_i(HH^{\dagger}) \right) \right], \tag{15}$$

where H is a $X_t \times X_r$ random matrix denoting the channel statistics. The quantity $\sigma_i(A)$ refers to the ith eigen-value of the matrix A where the eigenvalues are indexed in some arbitrary fashion. The interference $I(x,\phi)$ is just a scalar and is given by -

$$I(x,\phi) = \sum_{y \in \phi^T \setminus \{u\}} l(||y-x||), \tag{16}$$

where ϕ^T is the set of points on $\mathbf{S}$ corresponding to the transmitters and the transmitter of the receiver at location $x \in \phi$ is assumed to be present at $u \in \phi^T$. If one, employs the MIMO rate Equation (15), then we get the following result -

Corollary 3. *The critical arrival intensity of links λ_c under the foregoing assumptions is $\frac{CX_r}{La\ln(2)}$.*

We provide a proof sketch in Appendix XII. We phrase the above result as a corollary since it is not surprising to have as critical density in this case of independent MIMO channel with no CSIT as X_r times the critical density for a single-antenna link based SBD process. The total transmit power is 1 just as in the case with single antenna, however the presence of X_r receive antennas per link implies that the network can support upto X_r times more transmitters than in the case with single antenna link. This result indicates that the effect of having multiple antennas at the transmitter is not beneficial if there is no CSIT. On the other hand, in the presence of CSIT, one would expect to receive gain from the presence of multiple transmit antennas as S_x is a function of the channel realization H_{xx} thereby exploiting the diversity from multiple transmit antennas better. However, we do not pursue this question in the present paper and leave the analysis of the generalized MIMO system to future work.

VII. Conclusion and Future Work

In this paper, we proposed a novel space-time interacting particle system to model spectrum sharing in ad-hoc wireless networks. We computed exactly the phase-transition point for time ergodicity. We also proved the intuitive fact that the steady-state point-process corresponding to this dynamics exhibits clustering. In order to understand the performance metric of density of links in steady-state, we proposed a Poisson heuristic β_f (which is a bound subject to Conjecture 1) and a second order heuristic β_s. We saw from simulations that both the heuristics are tight in the two asymptotic regimes of heavy and light traffic. However, in the intermediate traffic regime, we found that the heuristic β_s performs much better compared to the Poisson heuristic β_f as β_s accounts for some spatial correlations which are non negligible in this regime. We also saw through simulations that any form of simplistic modeling of spatio-temporal interactions through PPP or equivalent queues ignoring spatial clustering, leads to poor estimates for performance.

From a mathematical perspective, we identified several challenging directions of future work in the simulation section. In particular, understanding the higher order moment measure of ϕ_0 will be key in evaluating or providing provably tighter bounds for performance metrics. Understanding the higher-order moment measures may also aid in making progress on Conjecture 1. From an information-theoretic perspective, we considered a dynamic interference network where links treat interference as noise. However, it will be interesting to consider other receiver schemes such as Successive Interference Cancellation or Joint-Decoding and show that the critical arrival rate for these schemes are strictly better than considering all Interference as Noise. This will then yield the complete dynamic version of the model considered in [17], namely a dynamic version of an interference network with point-to-point codes.

Acknowledgements

This work was supported by an award from the Simons Foundation (#197982) to The University of Texas at Austin. The authors also acknowledge the support of TACC (Texas Advanced Computing Center) for providing access to computing resources to perform the simulations.

References

[1] C. E. Shannon, "Two-way communication channels," in *Proc. of the 4th Berkeley Symp. on Mathematical Statistics and Probability, Vol 1*, Berkeley, Calif., 1961, pp. 611–644.

[2] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, 2009.

[3] F. Baccelli and B. Blaszczyszyn, *Stochastic geometry and wireless networks: Theory*. Now Publishers Inc, 2009, vol. 1.

[4] F. Baccelli, B. Blaszczyszyn, and P. Mühlethaler, "An aloha protocol for multihop mobile wireless networks," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 421–436, 2006.

[5] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.

[6] T. Bonald, S. Borst, N. Hegde, and A. Proutière, "Wireless data performance in multi-cell scenarios," *ACM SIGMETRICS Performance Evaluation Review*, vol. 32, no. 1, pp. 378–380, 2004.

[7] X. Lin, N. B. Shroff, and R. Srikant, "On the connection-level stability of congestion-controlled communication networks," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2317–2338, 2008.

[8] C. Moallemi and D. Shah, "On the flow-level dynamics of a packet-switched network," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 38, no. 1, 2010, pp. 83–94.

[9] F. P. Kelly, *Reversibility and Stochastic Networks*. New York, NY: Cambridge University Press, 2011.

[10] L. Massoulié and J. Roberts, "Bandwidth sharing: objectives and algorithms," in *IEEE INFOCOM*, vol. 3, 1999, pp. 1395–1403.

[11] C. Preston, "Spatial birth-and-death processes." Bull. Int. Stat. Inst., Proc. of the 40th Session Warsaw 46, No.2, 371-391 , 1975.

[12] N. L. Garcia and T. G. Kurtz, "Spatial birth and death processes as solutions of stochastic equations," *Alea*, vol. 1, pp. 281–303, 2006.

[13] M. D. Penrose, "Existence and spatial limit theorems for lattice and continuum particle systems," *Probab. Surveys*, vol. 5, pp. 1–36, 2008.

[14] F. Baccelli, F. Mathieu, I. Norros, and R. Varloot, "Can P2P networks be super-scalable?" in *IEEE INFOCOM*, 2013, pp. 1753–1761.

[15] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: volume II: General theory and structure*. Springer Science & Business Media, 2007, vol. 2.

[16] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[17] F. Baccelli, A. El Gamal, and D. N. Tse, "Interference networks with point-to-point codes," *Information Theory, IEEE Transactions on*, vol. 57, no. 5, pp. 2582–2596, 2011.

[18] M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic: Evidence and possible causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, 1997.

[19] R. W. Wolff, "Poisson arrivals see time averages," *Operations Research*, vol. 30, no. 2, pp. 223–231, 1982.

[20] J. Møller and R. P. Waagepetersen, "Modern statistics for spatial point processes," *Scandinavian Journal of Statistics*, vol. 34, no. 4, pp. 643–684, 2007.

[21] K. A. Hamdi, "A useful lemma for capacity analysis of fading interference channels," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 411–416, 2010.

[22] A. Baddeley and R. Turner, "spatstat: An R package for analyzing spatial point patterns," *Journal of Statistical Software*, vol. 12, no. 6, pp. 1–42, 2005.

[23] H. C. Gromoll, "Diffusion approximation for a processor sharing queue in heavy traffic," *The Annals of Applied Probability*, pp. 555–611, 2004.

[24] E. Telatar, "Capacity of multi-antenna gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[25] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[26] F. Baccelli and P. Brémaud, *Elements of queuing theory: Palm-martingale calculus and stochastic recurrence*. New York: Springer-Verlag, 1994.

[27] P. Robert, *Stochastic networks and queues*. Springer-Verlag, 2003.

[28] L. Massoulié, "Structural properties of proportional fairness: stability and insensitivity," *The Annals of Applied Probability*, pp. 809–839, 2007.

APPENDIX

VIII. PROOF OF THEOREM 1

Proof. We prove this by contradiction. Assume that ϕ_t is in stationary regime and that $\lambda > \frac{C}{\ln(2)La}$. We use the Miyazawa's Rate-Conservation Principle or Law (RCL) (e.g. [26], 1.3.3) to set-up a system of equations and identify a contradiction. Applying the RCL to the stochastic process $\phi_t(\mathbf{S})$ which counts the number of links yields,

$$\lambda|\mathbf{S}| = \lambda_d, \tag{17}$$

where λ_d is the intensity of the point-process on $\mathbb{R}$ corresponding to the epochs of a death-time. Since we assumed that ϕ_t is in stationary regime, the point process formed on the real line by the instants of a death is stationary with intensity $\lambda_d = \lambda|\mathbf{S}|$. Applying RCL to the total "work-load" in the network i.e. the total number of bits that each of the transmitters present are yet to send to their corresponding receivers, we get

$$\lambda|S|L = \mathbb{E}\left[\sum_{x\in\phi_0} R(x,\phi_0)\right], \tag{18}$$

where $R(x,\phi)$ is given in Equation (1). From the definition of Palm Probability of ϕ_0, we have that

$$\lambda|\mathbf{S}|L = \mathbb{E}^0_{\phi_0}\left[R(0,\phi_0)\right]\mathbb{E}[\phi_0(\mathbf{S})], \tag{19}$$

where $\mathbb{E}^0_{\phi_0}$ is the (spatial) Palm Probability of ϕ_0 and $\phi_0(\mathbf{S})$ is the random variable denoting the number of links in the network in steady-state. Note that from our assumption that ϕ_t is in stationary regime ensures the existence of the Palm Probability measure of the spatial point process ϕ_0. Applying rate-conservation to the stochastic process $\mathbf{I}_t = \sum_{x\in\phi_t} I(x,\phi_t)$, the sum interference seen at all receivers (which could possibly be ∞), we get

$$\lambda|\mathbf{S}|\mathbb{E}^{\uparrow}[\mathcal{I}] = \lambda_d\mathbb{E}^{\downarrow}[\mathcal{D}], \tag{20}$$

with $\mathcal{I} = \mathbf{I}_{0_+} - \mathbf{I}_0$ and $\mathcal{D} = \mathbf{I}_0 - \mathbf{I}_{0_+}$. Here, $\mathbb{E}^{\uparrow}$ denotes the (time) Palm probability corresponding to the point process on $\mathbb{R}$ of birth instants and $\mathbb{E}^{\downarrow}$ denotes the (time) Palm probability of the point process on $\mathbb{R}$ corresponding to the instants of death. From Equation (17) we have

$$\mathbb{E}^{\uparrow}[\mathcal{I}] = \mathbb{E}^{\downarrow}[\mathcal{D}]. \tag{21}$$

From the PASTA property and the fact that the births are uniform in $\mathbf{S}$, we have from Campbell's theorem that

$$\mathbb{E}^{\uparrow}[\mathcal{I}] = 2\mathbb{E}[\phi_0(\mathbf{S})]\frac{a}{|\mathbf{S}|}. \tag{22}$$

Since the file-sizes at all transmitters are i.i.d. exponential with mean L, the point process on the real line corresponding to the death-instants admits as stochastic-intensity $\mathbf{R}_t = \frac{1}{L}\sum_{x\in\phi_t} R(x,\phi_t)$ with respect to the filtration $\mathcal{F}_t = \sigma(\phi_s : s \le t)$, the sigma algebra corresponding to the locations. Hence, it then follows from Papangelou's theorem (e.g. [26], Theorem 1.9.2) that

$$\frac{d\mathbb{P}^{\downarrow}}{d\mathbb{P}}|_{\mathcal{F}_{0_-}} = \frac{\mathbf{R}_0}{\mathbb{E}[\mathbf{R}_0]}. \tag{23}$$

Since the decrease in total interference (in state ϕ_{0_-}) is of magnitude $I(X,\phi_0)$ with probability $\frac{R(X,\phi_0)}{L\mathbf{R}_0}$ if $X \in \phi_{0_-}$, we get

$$\begin{aligned}
\mathbb{E}^{\downarrow}[\mathcal{D}] &= 2\mathbb{E}\left[\frac{\mathbf{R}_0}{\mathbb{E}[\mathbf{R}_0]}\sum_{x\in\phi_0}\frac{R(x,\phi_0)}{L\mathbf{R}_0}I(x,\phi_0)\right] \\
&= 2\frac{\mathbb{E}[\sum_{x\in\phi_0} R(x,\phi_0)I(x,\phi_0)]}{L\mathbb{E}[\mathbf{R}_0]} \\
&= 2\frac{\mathbb{E}^0_{\phi_0}[R(0,\phi_0)I(0,\phi_0)]}{L\mathbb{E}[\mathbf{R}_0]}\mathbb{E}[\phi_0(\mathbf{S})].
\end{aligned} \tag{24}$$

Now combining, Equations (24), (22) and (18), we get

$$a = \frac{\mathbb{E}^0_{\phi_0}[R(0,\phi_0)I(0,\phi_0)]}{L\lambda}. \tag{25}$$

From Equation (1) and basic calculus, we have that $R(0,\phi_0)I(0,\phi_0) \leq \frac{C}{\ln(2)}$ which is a deterministic bound that is true for *any* $\phi \in \mathbf{M}(\mathbf{S})$. Applying this inequality to Equation (25), we get the inequality that

$$\lambda \leq \frac{C}{\ln(2)La}. \tag{26}$$

Inequality (26) is a contradiction to our assumption that ϕ_t is in stationary regime and that $\lambda > \frac{C}{\ln(2)La}$. □

IX. Proof of Theorem 2

For simplicity of the proof, we assume the link distance $T = 0$. The high level idea of the proof is that we tesselate the space $\mathbf{S}$ and study another "upper-bound" Markov Chain living on a countable state-space which we analyze through fluid limit techniques. We then conclude about the ergodicity of ϕ_t which is a Markov Chain on the topological space $\mathbf{M}(\mathbf{S})$.

To define the upper-bound chain, we first tessellate the square $\mathbf{S}$ into cells where each cell is a square of length exactly ϵ. Since $\mathbf{S}$ is a torus, we assume without loss of generality that the origin is in the center of a cell. One can find a sequence of such tessellations with the side length of the cells going to 0. The tessellation for each valid $\epsilon > 0$ results in n_ϵ, a finite number of cells as $\mathbf{S}$ is compact. Index the cells by i and let A_i denote the subset of $\mathbf{S}$ corresponding to cell i and $a_i \in A_i$ denote its center. The cell containing the origin is indexed 0 i.e. $a_0 = 0$. For such an ϵ tessellation, we define a new path-loss function $l_\epsilon(x,y)$ where $l_\epsilon(x,y) = l_\epsilon(a_i,a_j)$ for all $x \in A_i$ and $y \in A_j$ and

$$l_\epsilon(a_i,a_j) = \sup\{l(||b_i - b_j||) : ||b_i - a_i||, ||b_j - a_j|| \in \{0,\epsilon\}\}.$$

Note that the function l_ϵ satisfies

$$\sum_i l_\epsilon(a_i,a_j) = \sum_i l_\epsilon(a_i,0) = \frac{1}{\epsilon^2}\int_{x\in\mathbf{S}} l_\epsilon(||x||)dx, \tag{27}$$

since $\mathbf{S}$ is a square torus and each cell A_i is a square of side-length ϵ,

The upper bound Markov-Chain is denoted as $\phi^{(\epsilon)}_t$ which takes value in the space $\mathbf{M}(\mathbf{S})$. This chain has the exact same dynamics as described in Equation (4) *except* that the interference comes from $l_\epsilon(.,.)$ instead of from $l(\cdot)$,

Lemma 2. *For all time t, the point-process $\phi^{(\epsilon)}_t$ stochastically dominates ϕ_t. This implies that if $\phi^{(\epsilon)}_t$ is stable for a particular λ, then so is ϕ_t for that value of λ.*

Proof. We have from the monotonicity of $l(\cdot)$, $l_\epsilon(x,y) \geq l(x,y) = l(||x-y||)$ for each $x,y \in \mathbf{S}$. Thus, for each $x \in \mathbf{S}$ and each $\phi \in \mathbf{M}(\mathbf{S})$, $I_\epsilon(x,\phi) \geq I(x,\phi)$ and subsequently $R_\epsilon(x,\phi) \leq R(x,\phi)$ as $R(x,\phi)$ is a decreasing function of $I(x,\phi)$. Therefore the point process $\phi^{(\epsilon)}_t$ *stochastically dominates* the point process ϕ_t. This follows from the fact that for any $\phi \in \mathbf{M}(\mathbf{S})$, we have that the birth rate $\lambda|S|$ is the same for both process , whereas the death-rate of each point of $x \in \phi$ satisfies $\frac{1}{L}R_\epsilon(x,\phi) \leq \frac{1}{L}R(x,\phi)$. Also form Equation (1), if $\phi_1 \subseteq \phi_2$, then for each $x \in \phi_1 \cap \phi_2$, $R(x,\phi_1) \geq R(x,\phi_2)$. Hence one can construct a coupling of the process ϕ_t and $\phi^{(\epsilon)}_t$ such that $\phi_t \subseteq \phi^{(\epsilon)}_t$ $\forall t$, i.e. a point is alive in ϕ_t only if it is also alive in $\phi^{(\epsilon)}_t$. Therefore, if $\phi^{(\epsilon)}_t$ is ergodic for a given λ, then ϕ_t is also ergodic for that arrival rate λ. □

Define $\mathbf{X}^{(\epsilon)}(t) = \{\phi^{(\epsilon)}_t(A_i)\}_{i=1}^{n_\epsilon}$ as the n_ϵ dimensional vector taking values in $\mathbb{N}^{n_\epsilon}$. It is easy to see that $\mathbf{X}^{(\epsilon)}(t)$ is a Markov-Chain since the path-loss function $l_\epsilon(x,y)$ does not distinguish between two different locations of space inside a cell. It is also evident that if $\mathbf{X}^{(\epsilon)}(t)$ is ergodic, then $\phi^{(\epsilon)}_t$ is ergodic since $\lim_{t\to\infty}\mathbb{P}[\phi^{(\epsilon)}_t(\mathbf{S}) < \infty] = \mathbb{P}[||\mathbf{X}^{(\epsilon)}(t)||_1 < \infty] = 1$ where the last equality follows from the fact that $\mathbf{X}^{(\epsilon)}(t)$ is a finite-dimensional ergodic Markov chain on $\mathbb{N}^{n_\epsilon}$. Hence, a sufficient

condition for stability of ϕ_t is a condition for the Markov Chain $\mathbf{X}^{(\epsilon)}(t)$ to be ergodic.

We show in Theorem 6 that $\mathbf{X}^{(\epsilon)}(t)$ (and hence $\phi_t^{(\epsilon)}$) is ergodic if

$$\lambda < \frac{C}{L\ln(2)\int_{x\in\mathbf{S}} l^{\epsilon}(x,0)dx}, \tag{28}$$

which will actually conclude the proof of Theorem 2. This can be seen as follows. Since the point process ϕ_t^{ϵ} stochastically dominates ϕ_t, we can optimize the stability region in Equation (28) by choosing the best ϵ. As the function $r \to l(r)$ is monotone, $l_\epsilon(x,0)$ is monotone increasing in ϵ for each $x \in \mathbf{S}$ and hence we want to have ϵ as small as possible. Furthermore, the function $r \to l(r)$ has only a countable set of discontinuity points (as it is bounded non-increasing), we have that as ϵ goes to 0, $l_\epsilon(x,0)$ converges to $l(x,0)$ for almost-every $x \in \mathbf{S}$. Hence, $\lim_{\epsilon\to 0}\int_{x\in\mathbf{S}} l_\epsilon(||x||)dx = \int_{x\in\mathbf{S}} l(||x||)dx$ from the Monotone Convergence theorem. Therefore, if $\mathbf{X}^{(\epsilon)}(t)$ is ergodic under condition in Equation (28), then ϕ_t will be ergodic under the condition

$$\lambda < \limsup_{\epsilon\to 0} \frac{C}{L\ln(2)\int_{x\in\mathbf{S}} l^{\epsilon}(x,0)dx} = \frac{C}{L\ln(2)\int_{x\in\mathbf{S}} l(x,0)dx}, \tag{29}$$

which will conclude the proof of Theorem 2.

Theorem 6. *$\mathbf{X}^{(\epsilon)}(t)$ is ergodic under the condition in Equation (28).*

Proof. We can write the following evolution for the vector $\mathbf{X}^{(\epsilon)}(t)$ which we refer to as $\mathbf{X}(t)$ in the sequel for convenience as

$$\begin{aligned} &X_i \to X_i + 1 \text{ at rate } \lambda\epsilon^2 \\ &X_i \to X_i - 1 \text{ at rate} \\ &\quad X_i \log_2\left(1 + \frac{1}{N_0 + \sum_{j=1}^{n_\epsilon}(X_j - \mathbf{1}(j=i))l_\epsilon(a_i,a_j)}\right). \end{aligned}$$

Under condition in Equation (28), we show the following drift argument to hold which will conclude the proof.

Theorem 7. *[27] Let $\mathbf{X}(t)$ be a Markov Chain taking values in a countable state space $\mathcal{S}$. Assume there exists a function $L : \mathcal{S} \to \mathbb{R}_+$ and constants $A < \infty$, $\epsilon > 0$ and an integrable stopping time $\hat{\tau} > 0$ such that for all $x \in \mathcal{S}$:*

$$L(x) > A \implies \mathbb{E}_x L(\mathbf{X}(\hat{\tau})) \le L(x) - \epsilon\mathbb{E}_x(\hat{\tau}). \tag{30}$$

If in addition the set $\{x : L(x) \le A\}$ is finite and $\mathbb{E}_x L(\mathbf{X}(1)) < \infty$ for all $x \in \mathcal{S}$, then $\mathbf{X}(t)$ is ergodic.

We will show that the above theorem is satisfied with the Lyapunov function $L(x) = ||x||_\infty$ and

$$\hat{\tau} = L(\mathbf{X}(0))\left(\lambda\epsilon^2 - \frac{C}{L\ln(2)\sum_{k=0}^{n_\epsilon-1} l_\epsilon(a_k,0)}\right)^{-1} := L(\mathbf{X}(0))\tau, \tag{31}$$

a deterministic finite stopping-time. We will use the notation that $||x||_\infty = |x|$ which is also equal to $L(x)$.

To establish the drift condition, we pass to the fluid-limit. A fluid limit of the Markov-Process $\mathbf{X}(t)$ is denoted by $x(t)$ which is a n_ϵ dimensional vector. $x(t)$ is defined as a fluid limit if there exists non-decreasing Lipschitz continuous function $\{D_i(t)\}_{i=0}^{n_\epsilon-1}$ such that

$$x_i(t) = x_i(0) + \lambda\epsilon^2 t - D_i(t),$$

where the derivative of $D_i(t)$ satisfies
$\dot{D}_i(t) = \frac{Cx_i(t)}{L\ln(2)\sum_{j=0}^{n_\epsilon-1} x_j(t)l_\epsilon(a_i,a_j)}$, or equivalently, the fluid limit $x(t)$ satisfies the following set of differential equations. If $||x(t)||_\infty > 0$,

$$\frac{d}{dt}x_i(t) = \lambda\epsilon^2 - \frac{Cx_i(t)}{L\ln(2)\sum_{k=0}^{n_\epsilon-1} x_k(t)l_\epsilon(a_i,a_k)} \tag{32}$$

and if $||x(t)||_\infty = 0$,

$$\frac{d}{dt}x_i(t) = 0.$$

For $y \in \mathbb{R}^{n_\epsilon}$, denote by $S(y)$ the set of fluid functions $x(t)$ such that $x(0) = y$. The following theorem establishes that the above fluid equation is indeed obtained through an appropriate space and time scaling. It also establishes as a corollary that $S(y)$ is non-empty for any $y \in \mathbb{R}^{n_\epsilon}$.

Theorem 8. *Consider a sequence of initial conditions* $\{X^{(k)}(0)\}_{k\geq 1}$ *for the Markov Chain* $\mathbf{X}(t)$ *and a sequence of positive integers* $\{z_k\}_{k\geq 1}$ *with* $\lim_{k\to\infty} z_k = \infty$ *such that the limit* $\lim_{k\to\infty} z_k^{-1} X^{(k)}(0) = x(0)$ *exists. Then for all* $T > 0$ *and all* $\epsilon > 0$*, the following convergence takes place*

$$\lim_{k\to\infty} \mathbb{P}\left(\inf_{f\in S(x(0))} \sup_{t\in[0,T]} |z_k^{-1} X^{(k)}(z_k t) - x(t)| > \epsilon \right) = 0.$$

This proof is standard and is postponed later on in the appendix.

From the description of the dynamics, if we have $L(x(t)) = 0$, then $x_i(t) = 0$ for all i. Since $x(t)$ is a finite-dimensional vector, there exists at-least one coordinate $i^*(t)$ such that $x_{i^*(t)}(t) = L(x(t))$. Then one can write

$$\begin{aligned}\frac{d}{dt}L(x(t)) &= \lambda\epsilon^2 - \frac{CL(x(t))}{L\ln(2)\sum_{k=0}^{n_\epsilon-1} x_k(t) l_\epsilon(a_{i^*(t)}, a_k)} \\ &\leq \lambda\epsilon^2 - \frac{C}{L\ln(2)\sum_{k=0}^{n_\epsilon-1} l_\epsilon(a_k, 0)}, \end{aligned} \tag{33}$$

where the second inequality comes by the fact that $x_k(t) \leq L(x(t))$ and the symmetry of the torus as given in Equation (27). From Equation (33), we see that under the condition given in (28), $L(x(s)) = 0$ for all $s \geq \tau$ whenever $L(x(0)) = 1$, where $\tau = \left(\lambda\epsilon^2 - \frac{C}{L\ln(2)\sum_{k=0}^{n_\epsilon-1} l(a_k,0)}\right)^{-1}$, a deterministic time as defined in Equation (31).

Lemma 3. *If condition in Equation (28) holds, then*

$$\lim_{L(x)\to\infty} \frac{1}{L(x)} \mathbb{E}_x[|\mathbf{X}(L(x)\tau)|] = 0. \tag{34}$$

where τ *is defined in Equation (31)*

Proof. The first observation is that the family of random variables $\left\{\frac{|\mathbf{X}_x(|x|t)|}{|x|}\right\}_{x\in\mathbb{N}^{n_\epsilon}\setminus\{0\}}$ is uniformly integrable. Indeed, let $\{A_i(\cdot)\}_{i=0}^{n_\epsilon-1}$ be i.i.d. unit rate PPP denoting the arrivals into cell i. Then

$$X_i(t) \leq X_i(0) + A_i(\lambda\epsilon^2 t). \tag{35}$$

Thus, for $\mathbf{X}(0) = x$,

$$\frac{X_i(|x|\tau)}{|x|} \leq \frac{x_i}{|x|} + \frac{A_i(\lambda\epsilon^2|x|\tau)}{|x|}. \tag{36}$$

We have that $\frac{x_i}{|x|} \leq 1$ and the mean of $\frac{A_i(\lambda\epsilon^2|x|\tau)}{|x|}$ equal to $\lambda\epsilon^2\tau$. The variance of $\frac{A_i(\lambda\epsilon^2|x|\tau)}{|x|}$ is $\frac{\lambda\epsilon^2\tau}{|x|} \leq \lambda\epsilon^2\tau$ for all $x \in \mathbb{N}^{n_\epsilon} \setminus \{0\}$. As the variance is uniformly bounded, the random variables $\left\{\frac{X_i(|x|\tau)}{|x|}\right\}_{x\in\mathbf{N}^{N_\epsilon}\setminus\{0\}}$ are uniformly integrable. In addition, $\frac{|\mathbf{X}(|x|\tau)|}{|x|} \leq \sum_i \frac{X_i(|x|\tau)}{|x|}$, gives that $\left\{\frac{|\mathbf{X}_x(|x|\tau)|}{|x|}\right\}_{x\in\mathbf{N}^{N_\epsilon}\setminus\{0\}}$ is uniformly integrable since it is bounded above by a finite sum of random variables belonging to uniformly integrable families.

Let x_k be any sequence of initial conditions such that $|x_k| \to \infty$. This implies that $a_k = \mathbf{X}^{(k)}(0)/|x_k| = x_k/|x_k|$ with $a_k \in [-1,1]^{n_\epsilon}$ for all k. Since the cube $[-1,1]^{n_\epsilon}$ is compact, there is a convergent sub-sequence i.e. $\frac{\mathbf{X}^{k(l)}(0)}{|x_{k(l)}|} \to x(0)$ with $|x(0)| = 1$. From Theorem 8, there is a further sub-sequence of $k(l)$ such that $\frac{\mathbf{X}^{k'(l)}(|x_{k'(l)}|\tau)}{|x_{k'(l)}|} \to x(\tau)$ almost surely where the function $x(\cdot) \in S(x(0))$.

Under the stability condition (28), we have that for any fluid-limit function $x(\cdot) \in S(x(0))$, $x(\tau) = 0$ whenever $|x(0)| \leq 1$. This establishes that given any arbitrary sequence of initial conditions x_k with $|x_k| \to \infty$, one can find a further sub-sequence $k^{'}(l)$ such that

$$\lim_{k^{'}(l)\to\infty} \frac{1}{|x_{k^{'}(l)}|} |\mathbf{X}^{k^{'}(l)}(|x_{k^{'}(l)}|\tau)| = 0, \ a.s. \tag{37}$$

Therefore, we can conclude that for any sequence x_k with $|x_k| \to \infty$, we have $\frac{1}{|x_k|}|\mathbf{X}^k(|x_k|\tau)|$ tends to 0 in probability. But since, the family of random variables $\left\{\frac{|\mathbf{X}_x(|x|\tau)|}{|x|}\right\}_{x\in\mathbf{N}^{N_\epsilon}\setminus\{0\}}$ is uniformly integrable, we have that

$$\lim_{k\to\infty} \frac{1}{|x_k|} \mathbf{E}[|\mathbf{X}^k(|x_k|\tau)|] = 0. \tag{38}$$

As x_k was an arbitrary sequence, Equation (34) holds whenever condition (28) holds. □

From Equation (34), we have that for any $\epsilon > 0$, there is a large enough A_ϵ such that Equation (30) holds. Furthermore, for any finite A, the set $\{x \in \mathbf{N}^{N_\epsilon} : ||x||_\infty \leq A\}$ is finite. Hence, we have that $\mathbf{X}(t)$ is stable under the stability condition (28) which proves Theorem 2.

□

X. Proof of Theorem 3

Proof. The proof idea is to apply Rate-Conservation equations similar to that of Theorem 1. For any receiver-transmitter pair $(x; y) \in \phi_t$, define $B_t(x) = \sum_{T\in\phi_t^{tx}\setminus\{y\}} f(||T - x||)$ and the cadlag process $\mathcal{B}_t = \sum_{x\in\phi_t^{Rx}} B_t(x)$.

Since we assume that the dynamics ϕ_t is ergodic, we write RCL for the stochastic process $\mathcal{B}_t$

$$\lambda|S|\mathbb{E}\left[2\int_{x\in\mathbf{S}} B_0(x)\frac{dx}{|\mathbf{S}|}\right] = \lambda_d\mathbb{E}\left[\frac{\mathbf{R}_0}{\mathbb{E}[\mathbf{R}_0]}\sum_{T_n\in\phi_0}\frac{R(T_n,\phi_0)}{\mathbf{R}_0}2B_0(T_n)\right] \tag{39}$$

The LHS follows from PASTA and the fact that a birth can happen anywhere in $\mathbf{S}$ uniformly and independently. The RHS follows from the Papangelou's theorem that the point process on $\mathbb{R}$ corresponding to the death epochs admits $\mathbf{R}_t = \frac{1}{L}\sum_{X_n\in\phi_0} R(X_n, \phi_0)$ as its Stochastic Intensity with respect to the filtration $\mathcal{F}_t = \sigma(\{\sigma_s : s \leq t\})$, the sigma algebra generated by the location of the links. We also have $\lambda_d = \lambda|S|$ from Equation (17) and $\mathbb{E}[\mathbf{R}_0] = \lambda|S|L$ from Equation (18). Using this to simplify Equation (39), we get

$$\mathbb{E}[B_0(0)] = \frac{1}{\lambda|S|L}\mathbb{E}\left[\sum_{T_n\in\phi_0} R(T_n,\phi_0)B_0(T_n)\right], \tag{40}$$

where we used Fubini's theorem and the fact that ϕ_0 is stationary in simplifying the LHS. Using the definition of Palm probability to simplify the RHS, we get

$$\mathbb{E}[B_0(0)] = \frac{\beta|\mathbf{S}|}{\lambda L|\mathbf{S}|}\mathbb{E}^0_{\phi_0}\left[R(0,\phi_0)B_0(0)\right]. \tag{41}$$

Since both $f(\cdot)$ and the path-loss $l(\cdot)$ are positive non-increasing functions, we have the deterministic behavior that if $B_0(0)$ increases, then $R(0, \phi_0)$ decreases. Hence, we can use the association inequality

$$\mathbb{E}^0_{\phi_0}\left[R(0,\phi_0)B_0(0)\right] \leq \mathbb{E}^0_{\phi_0}\left[R(0,\phi_0)\right]\mathbb{E}^0_{\phi_0}\left[B_0(0)\right] \tag{42}$$

Employing Inequality (42) in Equation (41), and the RCL $\lambda L = \beta\mathbb{E}^0_{\phi_0}[R(0,\phi_0)]$ from equation (19) we get

$$\mathbb{E}[B_0(0)] \leq \mathbb{E}^0_{\phi_0}[B_0(0)]. \tag{43}$$

□

XI. Proof of Theorem 8

Proof. This can be argued by contradiction. Assume that for some $\epsilon > 0$ and a sub-sequence

$$\mathbb{P}\left(\inf_{f\in S(x(0)} \sup_{t\in[0,T]} |z_k^{-1}X(z_k t) - f(t)| > \epsilon\right) \geq \epsilon \tag{44}$$

Without loss of generality, assume the above holds true for all $k \geq 1$.

The trajectories of the process $X^k(t)$ can be written in terms of independent unit-rate Poisson process A_i^k and D_i^k

$$X_i^k(t) \quad = \quad X_i^k(0) \quad + \quad A_i^k(\lambda\epsilon^2 t) \quad - \quad D_i^k\left(\int_0^t X_i^k(u)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(t)}du\right)\right). \tag{45}$$

That is, $X^k(t)$ is a functional of the Point Process satisfying the set of Equations (45).

One can rewrite equation (45) by a change of variables as

$$\frac{1}{z_k}X_i^k(z_k t) \;=\; \frac{1}{z_k}X_i^k(0) + \frac{1}{z_k}A_i^k(\lambda\epsilon^2 z_k t) - \frac{1}{z_k}D_i^k\left(\int_0^{z_k t} X_i^k(u)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(t)}\right)du\right). \tag{46}$$

Now replacing u by $z_k l$, we get the following

$$\frac{1}{z_k}X_i^k(z_k t) = \frac{1}{z_k}X_i^k(0)+\frac{1}{z_k}A_i^k(\lambda\epsilon^2 z_k t)-\frac{1}{z_k}D_i^k\left(z_k\int_0^t X_i^k(z_k l)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(z_k l)}\right)dl\right), \tag{47}$$

which can be written as

$$\frac{1}{z_k}X_i^k(z_k t) \quad = \quad \frac{1}{z_k}X_i^k(0) \;+\; \lambda\epsilon^2 t \;-\; \int_0^t X_i^k(z_k l)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(z_k l)}\right)dl \;+\; \delta_i^k(t), \tag{48}$$

where the error term $\delta_i^k(t)$ satisfies the stochastic bound

$$\sup_{t\in[0,T]}|\delta_i^k(t)| \quad \leq \quad \frac{1}{z_k}\sup_{t\in[0,\lambda\epsilon^2 T]}|A_i^k(z_k t) \;-\; z_k t| \;+\; \frac{1}{z_k}\sup_{t\in[0,T\log_2(e)]}|D_i^k(z_k t) \;-\; z_k t|. \tag{49}$$

The error term $\delta_i^k(t)$ is bounded by the following lemma.

Lemma 4. *[28] Let Ξ be a unit rate PPP on the real line. Then for all $T > 0$ and all $\lambda > 0$,*

$$\mathbb{P}\left(\sup_{t\in[0,T]}|\Xi(t) - t| \geq \lambda T\right) \leq e^{-Th(\lambda)} + e^{-Th(-\lambda)} \tag{50}$$

where $h(\lambda) = (1+\lambda)\log(1+\lambda) - \lambda$.

This lemma in particular implies that there exists a sub-sequence $k(l), l \geq 1$ and a sequence $\epsilon(l) \to 0$ such that $\forall i$

$$\sum_{l\geq 1}\mathbb{P}\left(\sup_{t\in[0,T]}|\delta_i^{k(l)}(t)| \geq \epsilon(l)\right) < \infty$$

By Borel-Cantelli's lemma, there exists a sub sequence such that for all i, $\lim_{l\to\infty}\sup_{t\in[0,T]}|\delta_i^{k(l)}(t)| \to 0$ almost surely.

Now consider the random function $w_k(t) = \int_0^t X_{ij}^k(z_k l)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(z_k l)}\right)dl$ which is Lipschitz for each sample path ω, i.e.

$$w_k(t) - w_k(s) = \int_s^t X_{ij}^k(z_k l)\log_2\left(1+\frac{1}{N_0+I_i^\epsilon(z_k l)}\right)dl \tag{51}$$

$$\leq (t-s)\frac{\log_2(e)}{\sup_{x,y\in\mathbf{S}} l^\epsilon(x,y)} < \infty. \tag{52}$$

From the Arzela-Ascoli theorem, there exists a sub-sequence such that $w_k(t)$ converges uniformly on $[0,T]$ to a Lipschitz continuous function $D_i(t)$ for each sample path ω. This along with the bound on $\sup_{t\in[0,T]}|\delta_{ij}(t)|$ yields that there is a sub-sequence such that

$$\frac{X_i^k(z_k t)}{z_k} \to x_{ij}(t) := x_i(0) + \lambda_i t - D_i(t),\ a.s., \tag{53}$$

where the convergence happens uniformly over $[0,T]$. $D_i(t)$ is Lipschitz since $x_i(t)$ is Lipschitz continuous. It remains to show that $\dot{D}_i(t) = \frac{x_i(t)}{I_i(t)}$. Since $D_i(t)$ is Lipschitz continuous, by Rademacher's theorem, it is differentiable almost everywhere on $[0,T]$. For all $h>0$,

$$\int_t^{t+h} X_i^k(z_k l)\log_2\left(1+\frac{1}{N_0+I_i^{\epsilon}(z_k l)}\right)dl \to \int_t^{t+h}\frac{x_i(l)}{I_i^{\epsilon,f}(l)}dl.$$

This follows from dominated convergence and the Lipschitz continuity of $l \to x_{ij}(l)$. Therefore $\dot{D}_i(t) = \frac{x_i(t)}{I_i^{\epsilon,f}(t)}$.

Hence, we have shown that given any sequence of initial conditions $X^k(0)$ and number z_k such that the limit $\frac{X^k(0)}{z_k} = x(0)$ exists, we can find a sub-sequence k_l such that $\frac{X^{k_l}(z_{k_l}t)}{z_{k_l}}$ converges almost surely to the Lipschitz continuous fluid limit function $x(t)$. This is a contradiction and hence the theorem is proved. □

XII. Proof Sketch for Corollary 3

We just provide the proof outline for the necessary condition. The sufficient condition follows identically as in proof of Theorem 2. For necessary condition, note that all Rate-Conservation Equations hold (Equations 17, 18 and 24). In particular, we only have a different upper bound for $R(x;\phi_0)I(x;\phi_0)$ since the rate-function used is a different one. From Equation (15), we have

$$\begin{aligned}
\lambda L\int_{x\in\mathbf{S}} l(|x|)dx &\le \lim_{q\to\infty}\mathbb{E}_H\left[\sum_{i=1}^{X_r} q\log_2\left(1+\frac{1}{X_t(N_0+1q)}\sigma_i(HH^{\dagger})\right)\right]\\
&\overset{(a)}{=} \mathbb{E}_H\left[\sum_{i=1}^{X_r}\lim_{q\to\infty} q\log_2\left(1+\frac{1}{X_t(N_0+q)}\sigma_i(HH^{\dagger})\right)\right]\\
&= \frac{\log_2(e)}{X_t}\mathbb{E}_H\left[\sum_{i=1}^{X_r}\sigma_i(HH^{\dagger})\right]\\
&\overset{(b)}{=} \log_2(e)X_r,
\end{aligned} \tag{54}$$

where (a) follows from Monotone Convergence theorem and (b) follows from the fact that H is a matrix whose entries are i.i.d. complex-normal random variables with zero mean and unit variance. Re-arranging inequality (54) yields the necessary condition on the stability region for the MIMO channel model with independent channels.