Emergent Behavior in Multipartite Large Networks: Multi-virus Epidemics

Augusto Santos∗, José M. F. Moura♮, and João Xavier†

Abstract

Epidemics in large complete networks is well established. In contrast, we consider epidemics in non-complete networks. We establish the fluid limit macroscopic dynamics of a multi-virus spread over a multipartite network as the number of nodes at each partite or island grows large. The virus spread follows a peer-to-peer random rule of infection in line with the Harris contact process (refer to [1]). The model conforms to an SIS (susceptible-infected-susceptible) type, where a node is either infected or it is healthy and prone to be infected. The local (at node level) random infection model induces the emergence of structured dynamics at the macroscale. Namely, we prove that, as the multipartite network grows large, the normalized Markov jump vector process \( \left( \tilde{Y}_N^i(t) \right) \), collecting the fraction of infected nodes at each island \( i = 1, \ldots, M \), converges weakly (with respect to the Skorokhod topology on the space of càdlàg sample paths) to the solution of an \( M \)-dimensional vector nonlinear coupled ordinary differential equation. In the case of multi-virus diffusion with \( K \in \mathbb{N} \) distinct strains of virus, the Markov jump matrix process \( \left( Y_N^i(t) \right) \), stacking the fraction of nodes infected with virus type \( j, j = 1, \ldots, K \), at each island \( i = 1, \ldots, M \), converges weakly as well to the solution of a \( (K \times M) \)-dimensional vector differential equation that is also characterized.

Keywords: Emergent behavior, Epidemics, Multipartite interacting agents, Exact mean field.

∗ A. Santos is with the Dep. of Electrical and Computer Engineering, Carnegie Mellon University, USA, and Instituto de Sistemas e Robotica (ISR), Instituto Superior Técnico (IST), Av. Rovisco Pais, Lisboa, Portugal (augustos@andrew.cmu.edu).

♮ J. M. F. Moura is with the Dep. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA 15213, ph:(412)268-6341, fax: (412)268-3890 (moura@ece.cmu.edu).

† J. Xavier is with ISR, IST, Av. Rovisco Pais, Lisboa, Portugal (jxavier@isr.ist.utl.pt).

The work of José M. F. Moura and Augusto Santos was supported in part by the National Science Foundation under Grant # CCF–1011903 and in part by the Air Force Office of Scientific Research under Grant # FA–95501010291.

The work of J. Xavier and A. Santos was also supported by the Fundação para a Ciência e a Tecnologia (FCT) (Portuguese Foundation for Science and Technology) through the Carnegie Mellon|Portugal Program under Grant SFRH/BD/33516/2008, CMU-PT/SIA/0026/2009 and SFRH/BD/33518/2008, and by ISR/IST pluriannual funding (POSC program, FEDER).
I. INTRODUCTION

Many complex dynamical systems exhibit emergent behavior – a well-structured macroscopic dynamics induced by simple, possibly random, local rules of interacting agents. Flocks of birds, ant colonies, beehives, brain neural networks, invasive tumor growth, and epidemics are all examples of large scale interacting agents systems displaying complex adaptive functional behaviors. Under appropriate initial conditions, a flock of birds reaches consensus on its cruise velocity while each bird probes only its nearest neighbors dynamics without a preferred leader in the flock (refer to [2]). This gives rise to synchronized flocking flying formations. Ant colonies can design optimal trails to access sources of food even though no ant bears the cognitive ability to shape up the colony to its blueprint mature optimal global behavior. Roughly, each scout-ant wanders around randomly tracking the leftover pheromone released by its scout peers. Reference [3] establishes the emergent dynamics of an idealized stochastic network model for ant colonies as the fluid limit dynamics of the network model (as the colony grows large). Seizure is an intricate outcome of the complex neural network dynamics of the brain. Reference [4] presents an overview of graphical dynamical models that have been applied to better understand the nature of seizures and bridge the microscopical electrical activity in the brain with the clinical observations of the phenomenon.

The challenge in studying such large scale systems lies in their high-dimensionality plus the coupling among the agents via their interactions. Together, these are the needed ingredients to induce emergent behavior. For instance, consider \( N \) agents whose state-vector \( X^N(t) := (X_1(t), \ldots, X_N(t)) \) evolves as a jump Markov process over the state space \( S^N := \{0,1\}^N \). If the agents are independent, then it turns out that the state of each agent evolves as a jump Markov process and, moreover, any state construct \( (f(X_1(t), \ldots, X_N(t))) \), where \( f : \{0,1\}^N \to \mathbb{R}^M \) bears appropriate measurability properties (we skip the details here), is a Markov jump process. For instance, the fraction of agents at state 1, \( f(X_1(t), \ldots, X_N(t)) = \sum_{i=1}^N X_i(t)/N \), is Markov. Even for large \( N \), due to the independence assumption, a qualitative analysis of \( (X^N(t)) \) becomes tractable, but, in this example of independent agents, any weak law of large numbers will reflect the average behavior of each individual agent rather than an emergent global cooperative behavior. When the agents are coupled – e.g., an agent switches to state 1 with a rate that is proportional to the number of its neighbors in state 1–then, in general, neither the state of each agent is Markov nor the macroscopic low-dimensional states \( (f(X_1(t), \ldots, X_N(t))) \) are Markov and studying the microscopic high-dimensional dynamical system \( (X^N(t)) \) becomes quickly unfeasible with the number of agents \( N \). Establishing the emergent dynamics or, in other words, the functional weak law of large numbers under an arbitrary coupling topology of the agents is challenging.
For the special case of a complete topology of interaction – any agent evenly affects any other agent in the cloud, – low-dimensional macroscopic state-variables may still be Markov, even though the state of each individual agent is no longer Markov. Again, for complete networks, the fraction of infected nodes \( f(X_1(t), \ldots, X_N(t)) = \sum_{i=1}^{N} X_i(t)/N \) is Markov. Under this complete network setting, the emergent behavior is framed as the fluid limit dynamics of a global state variable \( (Y(t)) := (f(X_1(t), \ldots, X_N(t))) \) of interest. For example, reference [5] considers a multiclass flow of packets over a complete network with finite capacity nodes. It defines the macroscopic state variable \( (Y^N(t)) = (Y_1^N(t), \ldots, Y_L^N(t)) \) that collects the fraction of nodes \( Y_i^N(t) \) with a particular distribution \( i \) of packets over the different classes. Reference [5] proves that the empirical distribution \( (Y^N(t)) \) converges weakly, with respect to the Skorokhod topology on the space of sample paths, to the solution of a vector ordinary differential equation.

For general topologies, the evolution of macroscopic state variables is intricately tied to the high-dimensional microscopic state \( (X^N(t)) \) of the system. Reference [6] proposes to consider the impact of the topology on the diffusion of a virus in the network, but, to overcome the coupling difficulty that arises with non complete networks, reference [6] departs from a peer-to-peer diffusion model. The authors in [6] replace the exact transition rates of the microstate process \( (X(t)) \) by their average to establish their \( N \)-intertwined model. Were the states of the nodes independent processes (a very strong assumption) and the resulting \( N \)-intertwined model would be an exact model to describe the dynamics of the likelihood of infection of each node as pointed out by the authors.

In this paper, we go beyond the complete network model to establish the exact meanfield dynamics of a multi-virus epidemics over the class of multipartite networks, without making any artificial simplifying assumptions. We assume a stochastic network model for the peer-to-peer spread of different strains of virus among a cloud of agents to establish the emergent dynamics of the epidemics. The emergent behavior is the fluid limit dynamics of the fraction of infected nodes over time. Namely, we show that, when the number of agents goes to infinity in a way to be described momentarily, the fraction of infected agents at each island in the multipartite network converges weakly to the solution of a set of nonlinear ordinary differential equations.

We briefly outline the paper. In Section II, we set the problem formulation, defining the peer-to-peer stochastic network model underlying the microscopic dynamics of the diffusion of the strains of virus. In Section III, we establish the meanfield dynamics for a single virus spread over a bipartite network, that is, we prove that \( (\tilde{Y}^N(t)) = (\tilde{Y}_1^N(t), \tilde{Y}_2^N(t)) \) collecting the fraction of infected nodes at each island converges weakly to the solution of an ordinary differential equation. In Section IV, we extend
the meanfield proof in Section III to the general case of a multipartite network under multi-virus spread. Finally, in Section V, we conclude the paper.

II. PROBLEM FORMULATION

In this Section, we formally introduce the stochastic network process that models the peer-to-peer diffusion of multiple strains of virus over multipartite networks. First, we set up the environment where the epidemics takes place. Let \( G = (V, E) \) be an undirected network, where \( V = \{1, 2, \ldots, N\} \) and \( E = \{\{i, j\} : i, j \in V\} \) represent the set of nodes and edges of the graph \( G \), respectively. We say that two nodes \( i, j \in V \) are connected and represent it as \( i \sim j \), if \( \{i, j\} \in E \). In this paper, we establish the mean field dynamics of a multi-viral strain epidemics over the class of multipartite networks that is defined next.

**Definition 1 (Multipartite network)** A network \( G = (V, E) \) is multipartite if there exists a partition \( V = \{V_1, \ldots, V_M\} \) of \( V \) such that \( \{a, b\} \notin E \) for any \( a, b \in V_i \) and any \( i \in \{1, \ldots, M\} \). Also, \( u \in V_i, v \in V_j, u \sim v \Rightarrow w \sim r, \forall w \in V_i, r \in V_j \) with \( i \neq j \). When \( M = 2 \), the multipartite network is called **bipartite**.

The elements \( V_i \) of the partition \( V \) are referred to as *islands*. In words, if two nodes of different islands \( V_i \) and \( V_j \) are connected then any node from \( V_i \) is connected to any node from \( V_j \). In this case, we say that islands \( V_i \) and \( V_j \) are connected and refer to it as \( V_i \sim V_j \). This allows us to abstract the supertopology structure of islands as illustrated in Figure 1. Also, we refer to

\[ \mathcal{N}(V_i) = \{V_j : V_j \sim V_i\} \]

as the superneighborhood of island \( V_i \) and \( d_i = |\mathcal{N}(V_i)| \) refers to the superdegree or number of neighboring islands of island \( V_i \). As an example, for the superneighbors of island 1, in Figure 1, we have \( \mathcal{N}(1) = \{2, 5\} \) and thus, \( d_1 = 2 \).

Given a graph \( G = (E, V) \), we define the sequence of induced multipartite networks \( G^N = (E^N, V^N) \), indexed by \( N = (N_1, \ldots, N_M) \in \mathbb{R}^M \), where \( M \) is the fixed number of islands, \( N_i \) is the number of nodes at the \( i \)-th island of \( G^N \) and

\[ V_i^N \sim V_j^N \Leftrightarrow \ell_i \in V_i^N \sim \ell_j \in V_j^N \quad (1) \]

for all \( \ell_i, \ell_j \in V \), where \( V_i^N \) is the \( i \)-th island of \( G^N \), as depicted in figure 2. In words, all the multipartite graphs in the sequence \( G^N \) share the same supertopology imposed by the topology of \( G \), differing only
Figure 1: Multipartite network representation. Nodes from the same island cannot transmit the infection amongst themselves. Nodes from an island can transmit the virus to nodes in neighboring islands. For instance, any node from island 5 can infect any node from islands 1 and 2.

on the number of nodes per island that is given by the upper-index $N$. Given a graph $G$, we are interested in obtaining the limiting dynamics of the fraction of infected nodes per island and per strain over $G^N$ as $N$ grows to infinity.

Figure 2: Illustration of the multipartite network $G^N$, with $N = (3, 6, 5, 4, 8)$ and underlying graph $G$. 
Now, we introduce the dynamical model of diffusion. A node in a $G^N$ multipartite network may assume one of two possible states: infected or healthy and susceptible to infections. We define the binary tensor $X_{iIk}(t)$ as conveying the state of each node over time in $G^N$. By $X_{iIk}(t) = 1$, we refer to node $i$ at island $I$ being infected at time $t$, $t \geq 0$, with virus strain $k \in \{1, 2, \ldots, K\}$, and by $X_{iIk}(t) = 0$ if the corresponding node $i$ is healthy or infected with other strain $l \neq k$; in this latter case, $X_{iIl}(t) = 1$. The upper-index $N$ will permeate all relevant stochastic processes constructs to emphasize the underlying multipartite network $G^N$ induced by $G$. If only one strain of virus is present in the network then, for notational simplicity, we suppress the extra-index $k$ and rather write $X_{iI}(t) = 1$ to represent that node $i$ from island $I$ is infected at time $t$, $t \geq 0$, and $X_{iI}(t) = 0$ if otherwise.

Our microscopical model of diffusion of the virus is set at node level and goes as follows. If a node $i$ from island $I$ is $y$-infected—i.e., infected with the virus strain $y$—at time $t$, then it heals after an exponentially distributed random time $T^h_{iy}(i) \sim \text{Exp}(\mu^y_I)$ whose distribution depends on the island $I$ and on the type of infection $y$. Once a node $i$ in island $I$ is $y$-infected, it transmits the infection to a randomly chosen node at the neighbor island $J \in \mathcal{N}(I)$ after an exponentially distributed random time $T^c_{ijy}(i) \sim \text{Exp}(\gamma^y_{IJ})$ whose distribution only depends on the ordered pair $(I, J) \in \{1, 2, \ldots, M\}^2$ of communicating islands $I$ and $J$, and the type of infection $y$. Whenever considering only one strain of virus, we drop the strain subindex $y$. Also, if there is no room for ambiguity, we drop the node identity $i$, writing $T^h_{Iy}$ for the healing time and $T^c_{IJy}$ for the infection time of a node at island $I$. If the chosen node $j$ at island $J$ is already $k$-infected at the time of infection $t$, $t \geq 0$, then nothing happens; that is, $X^N_{jK}(t) = 1$. Therefore, $\sum_k X_{iIk}(t) \leq 1$ for all $i \in I$, $I \in \{1, \ldots, M\}$ and $t \geq 0$, or in words, a node can only be infected by one strain of virus at a time.

To summarize, an infected node $i \in I$ activates $d_I + 1$ independent exponentially distributed random variables, where $d_I = |\mathcal{N}(I)|$ of them are associated to the times for infection and one to the time for healing. Each of the $d_I$ random variables for infection is dedicated to one superneighbor $J \in \mathcal{N}(I)$ of island $I$ as illustrated in Figure 3. As an example, if a node from island 1 in Figure 3 is $y$-infected, then, after a time interval of length $T^c_{12y} \sim \text{Exp}(\gamma^y_{12})$, it picks randomly a node from island 2 and infects it. Also, after a time $T^c_{13y} \sim \text{Exp}(\gamma^y_{13})$, that is independent of $T^c_{12y}$ and the healing time $T^h_{1y}$, it chooses randomly a node from island 3 and infects it as long as it is still infected.

The microscopic process $(X^N(t))$ thus, evolves through jumps according to the triggering of a sequence of independent exponentially distributed random variables. All time service random variables are assumed to be independent and have support in a single probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $(\mathcal{F}^N_t)_{t \geq 0}$ the natural filtration induced by the sequence of independent random variables. That is, $\mathcal{F}^N_t$ gives us
Figure 3: Illustration of an infection. The infected (dark colored) node at island 1 activates three exponentially distributed random clocks, each dedicated to one neighbor island. The clock dedicated to island 2 rings after a time $T_{12}^c \sim \text{Exp}(\gamma_{12})$ and a node from island 2 is randomly picked to be infected. Also, the infected node in island 1 heals after a time $T_{1y}^h \sim \text{Exp}(\mu_1^y)$.

information on the values of all random variable times (healing or infection) involved in the evolution of $(X^N(t))$ up to time $t$, $t \geq 0$. Note that by construction $(X^N(t))$ is adapted to $(F^N_t)_{t \geq 0}$, i.e., $\sigma \{X^N(s) : 0 \leq s \leq t\} \subset F^N_t$, for all $t$, $t \geq 0$, where $\sigma \{X^N(s) : 0 \leq s \leq t\}$ represents the natural filtration of the process $(X^N(t))$.

Analyzing the full-microstate of the network over time according to the local infection model just presented becomes quickly unfeasible with the number of nodes in the network as the microscopic process $(X^N(t))$ is high-dimensional. Instead, we are interested in characterizing macroscopically the virus evolution in the multipartite network, namely, studying the dynamics of the number or fraction of infected nodes at each island. To fix ideas, we assume single virus epidemics for the rest of this Section, unless otherwise stated. We refer to $(Y^N(t)) = (Y^N(t) \cap 1, \ldots, Y^N(t) \cap M)$ as the macroprocess that stacks the number of infected nodes at each island. The normalized vector process $(\bar{Y}^N(t)) = (\bar{Y}^N_1(t), \ldots, \bar{Y}^N_M(t))$ collects the corresponding fractions of infected nodes per island, where $\bar{Y}^N_i(t) = Y^N_i(t)/N_i$ is the fraction of infected nodes at island $i$ at time $t$, $t \geq 0$, with $N_i = |V_i|$, the number of nodes at island $i$. The sequence of macroprocesses $(\bar{Y}^N(t))_N$ is indexed by $N = (N_1, \ldots, N_M)$, the vector collecting the cardinality $N_i$ of each island $i$ in the underlying multipartite network $G^N$. It turns out that, from the microscopic model of peer-to-peer infection previously described, $(Y^N(t))$ is a jump Markov process.
with transition rate matrix given by

\[
Q (\mathbf{Y}^N(t), \mathbf{Y}^N(t) - e_i) = \mu_i Y^N_i(t) \tag{2}
\]

\[
Q (\mathbf{Y}^N(t), \mathbf{Y}^N(t) + e_i) = \left( \sum_{j \sim i} \gamma_{ji} Y^N_j(t) \right) \left( \frac{N_i - Y^N_i(t)}{N_i} \right), \tag{3}
\]

where \( e_i \in \mathbb{R}^M \) is the canonical vector with the \( i \)th entry equal to 1 and the remaining entries equal to 0, and the lowercase subindexes in equations (2)-(3) refer to islands. In equation (2), we represent the rate to decrease the infected population at island \( i \) by one. This happens once any infected node from island \( i \) heals,

\[
T^h_i = \min \left\{ T^h_i(k) : X^N_{ki}(t) = 1, k \in i \right\} \sim \text{Exp} \left( \mu_i Y^N_i(t) \right).
\]

In equation (3), we represent the rate to increase by one the infected population at island \( i \). In this case, each neighboring island \( j \in N(i) \) of \( i \) will have \( Y^N_j(t) \) infected nodes and, thus, after a time

\[
T^c_i = \min \left\{ T^c_{ji}(k) : X^N_{jk}(t) = 1, k \in j, j \sim i \right\} \sim \text{Exp} \left( \sum_{j \sim i} \gamma_{ji} Y^N_j(t) \right), \tag{4}
\]

an attempt of infection will be made by a neighboring node at some neighboring island, where \( T^c_{ji}(k) \sim \text{Exp} \left( \gamma_{ji} \right) \) is the time that the infected node \( k \in j \) takes to make an attempt of infection towards a node at island \( i \). The \textit{minimum} in equation (4) runs over all infected nodes in all the neighboring islands of \( i \). The rate at which an infection from the neighboring islands takes to strike island \( i \) is, thus,

\[
\left( \sum_{j \sim i} \gamma_{ji} Y^N_j(t) \right).
\]

As referred in the microscopic model description, if an infection is transmitted to an already infected node, then the state of the sink node remains unchanged, that is, the infected population does not increase. Therefore, the effective rate of infection will be the rate at which infections arriving at island \( i \) hit a healthy node, that is, it is given by the arrival rate \( \left( \sum_{j \sim i} \gamma_{ji} Y^N_j(t) \right) \) times the probability of hitting a healthy node, which is equal to the fraction of healthy nodes at island \( i \) at time \( t \), \( \left( \frac{N_i - Y^N_i(t)}{N_i} \right) \), since the chosen \textit{victim} node is drawn uniformly randomly. Note that the topology of the underlying network impacts the increasing rate \( Q (\mathbf{Y}^N(t), \mathbf{Y}^N(t) + e_i) \) whereas the decreasing rate \( Q (\mathbf{Y}^N(t), \mathbf{Y}^N(t) - e_i) \) only relies on the number of infected nodes at the network at time \( t \), regardless of the peer-to-peer connections. Note that two or more events–infection or healing of a node–happens at the same time with probability zero, that is, the evolution of the vector macroprocess \( (\mathbf{Y}^N(t)) \) is driven almost surely through unit jumps of \( e_i \) at each time. The goal of the next Section is to explore the Markov structure of the macroprocess \( (\mathbf{Y}^N(t)) \) to establish weak convergence as the number of nodes per island \( N \) goes to infinite keeping the underlying graph \( G \) (and thus, the number of islands \( M \)) fixed. Namely, the
normalized process \( \mathbf{Y}^N(t) \) admits a decomposition in terms of a martingale and a drift term that is a functional built upon the transition rates previously presented,

\[
\mathbf{Y}^N(t) = \mathbf{Y}^N(0) + \left( \sum_{\text{martingale}} + \int_{0}^{t} \left( \sum_{\text{drift}} F(\mathbf{Y}^N(s-)) \right) ds \right). 
\]

This decomposition is also known as Dynkin’s formula (refer to [7]) and holds for any Markov process. In the next Section, we provide an explicit characterization of the martingale term, and we will establish that it converges weakly to zero as the underlying \( G^N \) multipartite network grows large, and as a result (to be also proved) the vector process will converge weakly to the solution of the deterministic ordinary differential equation

\[
\frac{d}{dt} \mathbf{Y}(t) = F(\mathbf{Y}(t)),
\]

where the vector field \( F \) will be characterized momentarily.

### III. MEAN FIELD – BIPARTITE SINGLE VIRUS

In this Section, we establish for a single virus spread over a bipartite network that the empirical distribution sequence \( \left( \mathbf{Y}^N(t) \right) \) converges weakly, as the network grows large, to the solution of a deterministic vector differential equation. By the network grows large, we mean that \( N_m \to \infty \) for all \( m \in \{1, \ldots, M\} \) with a finite asymptotic ratio \( N_i/N_j \to \alpha_{ji} < \infty \) between neighboring island sizes, as will be clearer momentarily, keeping the number of islands \( M \) fixed. To fix ideas, we consider throughout this Section a single-virus epidemics in a bipartite network. We extend the analysis to the multivirus epidemics over a general multipartite network in Section IV. We remark that the Markov jump process \( \left( \mathbf{Y}^N(t) \right) \) admits a decomposition into a martingale term plus a drift term obtained from transition rates as characterized in the previous Section. In Section III-A, we fix needed notation. Section III-B provides a pathwise characterization for the process \( \left( \mathbf{Y}^N(t) \right) \), in particular, the pathwise description of the martingale term \( \left( \mathbf{M}^N(t) \right) \) that will be later important to establish the weak convergence to zero of its normalized counterpart \( \left( \mathbf{M}^N(t) \right) = (M_1^N(t)/N_1, \ldots, M_M^N(t)/N_M) \). To prove the weak convergence of the normalized process \( \left( \mathbf{Y}^N(t) \right) \), we start by showing in Section III-C that the underlying sequence of martingales \( \left( \mathbf{M}^N(t) \right) \) converges weakly (with respect to the Skorokhod topology) to zero as \( N_i \) grows large for all \( i = 1, \ldots, M \) with \( M \in \mathbb{N} \), the number of islands, kept fixed. In Section III-D, we show that as a consequence \( \left( \mathbf{Y}^N(t) \right) \) is a tight family indexed by \( N \) whose accumulation points (for the weak convergence) are necessarily given by the solutions of a differential equation. By uniqueness of the resulting differential equation (the vector field is globally-Lipschitz), any convergent subsequence
converges to the solution of this differential equation. Therefore, since the limiting process is unique, it follows that the whole sequence converges to the solution of the differential equation.

A. Preliminary Notation

We briefly present the notation used throughout this Section.

- \( Y_N^i(t) \): number of infected nodes at island \( i \) at time \( t, t \geq 0 \). The boldface upperscript \( N = (N_1, N_2, \ldots, N_M) \) stands for a vector stacking the number of nodes \( N_i \) at each island \( i \). For the purpose of this Section \( M = 2 \).
- \( Y_N^i(t) \): fraction of infected nodes at island \( i \) at time \( t, t \geq 0 \), \( Y_N^i(t) = Y_N^i(t)/N_i \).
- \( Y_N^i(t) \): vector stacking the number of infected nodes at each island \( i \) at time \( t, t \geq 0 \), \( Y_N(t) = (Y_1^N, \ldots, Y_M^N) \).
- \( Y_N^i(t) \): vector stacking the fraction of infected nodes at each island \( i \) at time \( t, t \geq 0 \), \( Y_N(t) = (Y_1^N, \ldots, Y_M^N) \).
- \( N_\alpha: \mathcal{B}[0,\infty) \to \mathbb{N} \): Poisson point process with rate \( \alpha, \alpha > 0 \). \( N_\alpha(A) \) counts the number of events in the Borel-set \( A \in \mathcal{B}[0,\infty) \). We index the elements of a family of independent Poisson point processes by an upperscript \( N_{\alpha}^{(i)} \). More details are presented in the next Subsection III-B.
- \( Y_N^i \subset \mathbb{N}^M \): state-space of the process \( (Y_N(t)) \), defined by
  \[
  Y_N^i = \left\{ Y = (Y_1, \ldots, Y_K) \in \mathbb{N}^K : 0 \leq \sum_{k=1}^K Y_k \leq N_i \right\}, \forall i = 1, \ldots, M. \tag{5}
  \]
- \( D_\mathbb{R}[0,T] \): space of càdlàg (continue à droite, limite à gauche) functions \( f: [0,T] \to \mathbb{R} \) endowed with the Skorokhod topology, e.g., [8].
- \( \mathcal{B}(\mathbb{R}^n) \): Borel \( \sigma \)-algebra over \( \mathbb{R}^n \) with the standard topology.
- \( \text{Leb}() : \mathcal{B}(\mathbb{R}) \to [0,\infty] \): Lebesgue measure over the real line \( \mathbb{R} \).
- \( \left( Y_N(t) \right) \Rightarrow \left( \overline{Y}(t) \right) \): stands for \( \left( Y_N(t) \right) \) converges weakly, in the Skorokhod topology, to the process \( \left( \overline{Y}(t) \right) \).

B. Pathwise Representation

In this Section, we provide a pathwise characterization for the macroprocess \( (Y_N(t)) \) built upon the microscopic diffusion model. We briefly present the relevant definitions regarding point processes over the real line \( \mathbb{R} \) that will be the building blocks for the pathwise description of \( (Y_N(t)) \). For more details, refer to [8].
Definition 2 (Point measure) $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ is a point measure on $\mathbb{R}$ if there exists a sequence $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, so that

$$\mu(A) = \#(A \cap \{a_n\}_{n=1}^{\infty}) = \sum_{i=1}^{\infty} 1_{\{a_i \in A\}}, \quad \forall A \in \mathcal{B}(\mathbb{R}),$$

that is, $\mu(A)$ counts the number of points of the sequence $a_n$ in $A$ for any Borelian set $A \in \mathcal{B}(\mathbb{R})$.

We represent the set of point measures on $\mathbb{R}$ as $M_p(\mathbb{R})$. Therefore, to each point measure $\mu \in M_p(\mathbb{R})$ there exists an underlying real sequence $(a_n)_{n \in \mathbb{N}}$. A point measure $\mu$ on $\mathbb{R}$ is called Radon, if each compact interval entails only a finite number of elements of the associated sequence, or in other words the set of accumulation points of $(a_n)$ is empty. We now define a point process on the real line $\mathbb{R}$.

Definition 3 (Point process) $\mathcal{N}$ is a point process if it is a Radon point measure valued random variable,

$$\mathcal{N} : \Omega \to M_p(\mathbb{R})$$

$$\omega \to \mathcal{N}(\omega, \cdot).$$

In words, each realization $\omega \in \Omega$ leads to a sequence of points in the real line (void of accumulation points) that underlies the point measure $\mathcal{N}(\omega, \cdot)$. For simplicity, we refer to the random measure that counts the number of events in a Borel set as $\mathcal{N}(\cdot)$. Thus, $\mathcal{N}(A)$ is the random measure of $A \in \mathcal{B}(\mathbb{R})$.

We now introduce the definition of Poisson point process that will be central to building the macroscopic process $(Y^N(t))$ from the local (at node level) rules of infection.

Definition 4 (Poisson point process) $\mathcal{N}_\gamma$ is a Poisson point process on $\mathbb{R}$ with rate $\gamma > 0$ if it satisfies the following two conditions

1) **[independent increments]** Given $I_1, I_2, \ldots, I_n \subset \mathbb{R}$, $n$ disjoint intervals in the real line then, $\mathcal{N}_\gamma(I_1), \mathcal{N}_\gamma(I_2), \ldots, \mathcal{N}_\gamma(I_n)$ are independent random variables.

2) **[increment stationarity]** Let $I_1, I_2 \subset \mathbb{R}$ be two intervals. Then,

$$L := \text{Leb}(I_1) = \text{Leb}(I_2) \Rightarrow \mathcal{N}_\gamma(I_1) \overset{d}{\sim} \mathcal{N}_\gamma(I_2),$$

that is, $\mathcal{N}_\gamma(I_1)$ and $\mathcal{N}_\gamma(I_2)$ are Poisson distributed random variables with rate (or intensity) parameter $\gamma L$.

Definition 4 implies that the interarrival time interval between the elements of the underlying random sequence of the Poisson point process is exponentially distributed with mean $1/\gamma$ (refer to [9]). That is, the random sequence of points underlying the point process is constructed so that the time between
events $T = a_{n+1} - a_n \sim \text{Exp}(\gamma)$ is exponentially distributed. We may refer to a Poisson point process $\mathcal{N}_\gamma$ with rate $\gamma > 0$ as $\gamma$-Poisson process.

If a process $Z(t)$ counts the number of events up to time $t$, $t \geq 0$, from a Poisson source with rate $\gamma$, then,

$$Z(t) = \mathcal{N}_\gamma ((0, t]) .$$

As a concrete example, consider a permanently infected node at island 1 from Figure 3. Then, according to the nearest-neighbor infection model described at the end of Section II, $Z(t)$ counts the number of infections that arrives at island 2 due to this permanently infected source up to time $t$, $t \geq 0$, if we consider $\gamma_{12} = \gamma$. If instead of a single permanently infected node we had two, then the rate would double and the process counting the number of infections arriving at island 2 would still be Poisson given by

$$Z(t) = \mathcal{N}_{2\gamma} ((0, t]) .$$

If the process $(Z(t))$ counts the number of events from a discrete-time varying Poisson source, that is, if the real line $\mathbb{R}$ can be partitioned into intervals $\bigcup_{i=1}^{\infty} I_i = \mathbb{R}$ so that at each time interval $I_i$ there is a $\gamma_i$-Poisson source $\mathcal{N}^{(i)}$ acting then

$$Z(t) = \sum_{i=1}^{\infty} \mathcal{N}^{(i)}_{\gamma_i} (I_i \cap (0, t]) . \quad (6)$$

From equation (6) and the fact that $\mathcal{N}^{(i)}_{\gamma_i}$ are random measures, we have the following integral characterization for the process $(Z(t))$

$$Z(t) = \sum_{i : I_i \cap (0, t] \neq \emptyset} \int_0^t \mathbf{1}_{\{s \in I_i\}} \mathcal{N}^{(i)}_{\gamma_i} (ds) ,$$

where the integrals are taken with respect to the respective Poisson random measures $\mathcal{N}^{(i)}_{\gamma_i}$. That is, for each realization $\omega \in \Omega$, $\mathcal{N}^{(i)}_{\gamma_i} (\omega, \cdot)$ is a measure over $\mathbb{R}$, and that is the measure under which the integral is defined. For the virus spread case, we can consider that now the number of infected nodes at island 1 changes over time (instead of the static example with permanently infected sources). In this case, the process $(Z(t))$ counting the number of infections arriving at island 2 due to infected nodes in island 1 is given by

$$Z(t) = \sum_{k=1}^{N_1} \int_0^t \mathbf{1}_{\{Y_{n-\gamma} = k\}} \mathcal{N}^{(k)}_{\gamma} (ds) , \quad (7)$$
where, in this case, \( N^{(1)}, N^{(2)}, \ldots, N^{(N_1)} \) are independent Poisson processes. That is, we partitioned the real line up to time \( t \), \( (0, t] = \sum_{k=1}^{N_1} \{0 < s \leq t : Y^N_1(s) = k\} \) according to the number of infected nodes \( Y^N_1(s) = k \) at island 1 during the time interval \((0, t]\). During the time intervals where island 1 has \( k \) infected nodes, \( Y^N_1(s) = k \), the source of infection that strikes island 2 is Poisson with rate \( k\gamma \).

Equation (7) represents a sample path characterization for the process \( (Z(t)) \) that counts the number of infections that flow from island 1 towards island 2 up to time \( t, t \geq 0 \). Also, if each event \( a_n \) from a \( \gamma \)-Poisson source is only counted with probability \( p \) then,

\[
Z(t) = N_{\gamma p}((0, t]) .
\]

Namely, in this case \((Z(t))\) counts the number of infections from a permanently infected node from island 1 to 2 that strikes healthy nodes, assuming that a fraction of \((1 - p)\) of the nodes at the sink island 2 are always infected.

We now consider all these effects together to build the sample path characterization of our macroprocess \((Y^N(t))\). As mentioned, in this Section we look at the bipartite network single virus case. Let \((I_1(t))\) and \((H_1(t))\) be the processes counting the number of nodes that are or were infected (at least once) and number of healings, respectively, up to time \( t, t \geq 0 \), at island 1. We have

\[
Y^N_1(t) - Y^N_1(0) = I_1(t) - H_1(t)
\]

\[
= \sum_{\ell=1}^{N_1} \sum_{q=1}^{N_2} N^{(\ell,q)}(\frac{N_1-\ell}{N_1})(\{0 \leq s \leq t : Y^N_1(s) = \ell, Y^N_2(s) = q\})
\]

\[
- \sum_{\ell=1}^{N_1} N^{(\ell)}(\{0 \leq s \leq t : Y^N_1(s) = \ell\})
\]

\[
= \sum_{\ell=1}^{N_1} \sum_{q=1}^{N_2} \int_0^t 1_{\{Y^N(s) = \ell, Y^N(s) = q\}} \left( N^{(\ell,q)}(\frac{N_1-\ell}{N_1})(ds) \right)
\]

\[
- \sum_{\ell=1}^{N_1} \int_0^t 1_{\{Y^N(s) = \ell\}} \left( N^{(\ell)}(ds) \right)
\]

Note that the difference between the process \((I_1(t))\) and the process \((Z(t))\) in equation (7) is that \((Z(t))\) counts the number of arrival infections at island 1, which might also hit already infected nodes, whereas \((I_1(t))\) counts the number of effective infections that hits healthy nodes and thus increases the infected population. In the latter case, one has to account for the effect described in equation (8), where only a fraction \(\frac{N_1-k_1}{N_1}\) of nodes at island 1 is healthy at time \(t\), if \(Y^N_1(t) = k_1\). For a general multipartite

May 7, 2014 DRAFT
network, the pathwise dynamics is given by

$$Y_{i}^{N}(t) = Y_{i}^{N}(0) + \sum_{q=1}^{N_{i}} \sum_{\ell=1}^{N_{i}} \int_{0}^{t} \sum_{j \sim i} 1 \{ Y_{j}^{N}(s-)=q, Y_{i}^{N}(s-)=\ell \} \left( \mathcal{N}^{(\ell,q)}_{\gamma_{j,q}\left(\frac{n_{i}-\ell}{n_{i}}\right)}(ds) \right)$$

where all Poisson point processes $\mathcal{N}^{(\ell,q)}_{\gamma_{j,q}\left(\frac{n_{i}-\ell}{n_{i}}\right)}$ and $\mathcal{N}^{(\ell)}_{\ell \mu_{i}}$ indexed by $\ell$ and $q$ are independent. This is an important fact from the peer-to-peer model that will be evoked latter in this Section and in Section IV. Also, for notational simplicity, we drop from now on the upper-indexes of the Poisson processes. Observe that the inter-transmission term is the one that relies on the supertopology of the multipartite network.

Next, we frame the normalized martingale term $(\overline{M}^{N}(t))$ hidden within the pathwise characterization of the normalized process $(\overline{Y}^{N}(t))$ and prove that it converges weakly to zero as the network $G^{N}$ grows large. This loosely implies that the randomness of the process $(\overline{Y}^{N}(t))$ dies out as the number of agents grows.

C. Martingale Vanishes

We start by characterizing the martingale term $(\overline{M}^{N}(t))$ of our macroprocess $(\overline{Y}^{N}(t))$

$$\overline{Y}^{N}(t) = \overline{Y}^{N}(0) + \overline{M}^{N}(t) + \int_{0}^{t} F(\overline{Y}^{N}(s-)) \, ds$$

and afterwards we explore its structure to prove that its normalized counterpart

$$\left( \overline{M}^{N}(t) \right) := \left( M_{1}^{N}(t)/N_{1}, M_{2}^{N}(t)/N_{2} \right)$$

converges weakly to 0 as the number of nodes at each island goes to infinite with the number of islands $M = 2$ kept fixed. In words, this means that the randomness of the normalized macroprocess $(\overline{Y}^{N}(t))$ dies out as the network grows large. We start by showing that $\overline{M}^{N}(t)$ converges to zero in $L_{2}$ for all $t, t \geq 0$. Then, by Doob’s inequality, this will imply that it converges in probability, in the Skorokhod space of càdlàg sample paths, to zero, as will be clearer momentarily. Finally, this implies that the martingale converges weakly to zero. As for the rest of Section III, we concentrate on the case of single virus spread over a bipartite network. The stochastic vector process $(\overline{Y}^{N}(t))$ over the bipartite
network admits the following pathwise characterization:

\[
Y_i^{N}(t) = Y_i^{N}(0) + \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_i} \int_0^t 1\{Y_j^{N}(s-)=\ell, Y_j^{N}(s-)=q\} \left( \frac{N^{r(\ell,q)}_{\gamma,\ell,q} \gamma_{\ell,q}}{N_{\gamma}} \right) (ds) - \sum_{\ell=1}^{N_i} \int_0^t 1\{Y_i^{N}(s-)=\ell\} \left( \frac{N^{r(\ell)}_{\ell,\mu_i}}{N_i} \right) (ds),
\]

for \( i, j \in \{1, 2\} \) and \( i \neq j \). The inter-transmission term in equation (13) accounts for the number of infections transmitted from island \( j \) to healthy nodes in island \( i \) up to time \( t \), \( t \geq 0 \). The healing term accounts for the number of healings that occur in island \( i \) during the time interval \([0, t]\). One can check that almost surely the normalized process \( \tilde{Y}_i^{N}(t) = Y_i^{N}(t)/N_i \in [0, 1] \), \( \forall t \geq 0 \) and all \( i = 1, 2 \), if \( \tilde{Y}_i^{N}(0) \in [0, 1] \) a.s., that is, the set \([0, 1] \times [0, 1]\) is invariant for the stochastic dynamics of \( \left( \tilde{Y}_i^{N}(t) \right) \) (refer to [10]). In words, the underlying stochastic dynamics are consistent with our intuition about the underlying meaning of \( \left( \tilde{Y}_i^{N}(t) \right) \) that it is the fraction of infected nodes at island \( i \), and so it is clearly a quantity between 0 and 1, for all \( t \), \( t \geq 0 \). The next Theorem states that equation (13) can be further decomposed as equation (14) into a martingale \( \left( \tilde{M}_i^{N}(t) \right) \) plus a drift term, with an explicit characterization for the martingale term provided in the proof.

**Theorem 5 (Process Decomposition)** Let \( \left( Y_i^{N}(t) \right) \) be the number of infected nodes at island \( i \) at time \( t \) for a bipartite network with two islands 1 and 2. Let \( \gamma_{ji} \) be the rate at which a node from island \( j \) attempts to infect a node in island \( i \). Then, \( \left( Y_i^{N}(t) \right) \) admits the following pathwise characterization:

\[
Y_i^{N}(t) = Y_i^{N}(0) + M_i^{N}(t) + \int_0^t \gamma_{ji} Y_j^{N}(s-) \left( \frac{N_i - Y_i^{N}(s-)}{N_i} \right) ds - \int_0^t \mu_i Y_i^{N}(s-) ds,
\]

for \( i, j = 1, 2 \), with \( i \neq j \), where \( \left( M_i^{N}(t) \right) \) is a martingale and \( N_i \) is the number of nodes in island \( i \).

**Proof:** This follows as a Corollary to Dynkin’s Lemma (see [7]), but we will provide an explicit characterization for the martingale \( \left( M_1^{N}(t), M_2^{N}(t) \right) \) resulting from the microscopical diffusion model. It is easy to check that the well-known compensated Poisson process

\[
M(t) = N_{\gamma}(\left(0, t]\right) - \gamma t
\]
is a martingale adapted to the natural filtration \( \Sigma_t \coloneqq \sigma \{ \mathcal{N}_i(s) : 0 \leq s \leq t \} \) of \((\mathcal{N}_i(t))\). Now, by compensating the Poisson point processes involved in equations (13), we obtain

\[
Y_i^N(t) = Y_i^N(0) + \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_i} \int_0^t \mathbb{1}_{Y_j^N(s-) = q, Y_i^N(s-) = \ell} \left( \mathcal{N}_i^{(\ell,q)}(\lambda q_i N_i) (ds) - \gamma_{ji} q \left( \frac{N_i - \ell}{N_i} \right) ds \right) \\
- \sum_{k=1}^{N_i} \int_0^t \mathbb{1}_{Y_i^N(s-) = k} (N_k \mu_i (ds) - k \mu_i ds) \\
+ \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_i} \int_0^t \mathbb{1}_{Y_j^N(s-) = q, Y_i^N(s-) = \ell} \gamma_{ji} q \left( \frac{N_i - \ell}{N_i} \right) ds - \sum_{k=1}^{N_i} \int_0^t \mathbb{1}_{Y_i^N(s-) = k} k \mu_i ds.
\] (15)

It turns out that the martingale property of a compensated Poisson process is stable under the integration of a predictable process, [8], in particular, if \((Y(t))\) is a càdlàg adapted process, then

\[
M(t) = \int_0^t Y(s-) (\mathcal{N}_\lambda(ds) - \lambda ds)
\]

is a martingale adapted to the natural filtration \( \sigma \{ \mathcal{N}_\lambda(s) : 0 \leq s \leq t \} \) (refer to [11]). Moreover, the space \( \mathcal{M} \) of martingales adapted to the same filtration conforms a vector space and, therefore, \( M_{ai}^N(t) \) and \( M_{bi}^N(t) \) are martingales and

\[
M_i^N(t) = M_{ai}^N(t) - M_{bi}^N(t)
\] (16)

is a martingale. Therefore, the terms \( M_{ai}^N(t) \) and \( M_{bi}^N(t) \) in equation (15) are martingales with respect to the natural filtration of the underlying Poisson point processes \( \mathcal{N}_i^{(\ell,q)}(\lambda q_i N_i) \) and \( \mathcal{N}_\ell^{(\ell,q)} \), respectively, for \( \ell \leq N_i \) and \( q \leq N_j \). Theorem 5 is completed and the martingale term \((M_i^N(t))\) is characterized.

Next, we prove that the variance of the normalized zero-mean martingales \( \overline{M}_i^N(t) = (M_i^N(t)/N_i) \) converge to zero (as \( N \) grows large) for all \( t, t \geq 0 \), and \( i = 1, 2 \), that is \( \overline{M}_i^N(t) \) converges to zero in \( L_2 \), for all time \( t \).

**Theorem 6** Let \( \overline{M}_i^N(t) = (M_{ai}^N(t) - M_{bi}^N(t)) \) be described as in equation (15). We have

\[
\text{Var} \left( \overline{M}_i^N(t) \right) = E \left( \overline{M}_i^N(t) \right)^2 \longrightarrow 0, \quad \forall t \geq 0 \text{ and } i = 1, 2
\]

as \( N_i \to \infty \) for \( i = 1, 2 \) with \( N_i/N_j \to \alpha_{ji} < \infty \).

**Proof:** Note first that for \( s < t \)

\[
E \left( M_i^N(t) \right) = E \left( E \left( M_i^N(t) | \mathcal{F}_s \right) \right) = E \left( M_i^N(s) \right).
\]
So, \((M_i^N(t))\) is zero-mean, \(E(M_i^N(t)) = E(M_i^N(0)) = 0\) for all \(t \geq 0\); thus, \(\text{Var}(M_i^N(t)) = E(M_i^N(t)^2)\). Before proceeding, we state that orthogonality of compensated Poisson martingales is preserved under integration of predictable processes, which will be important to complete the proof of Theorem 6.

**Theorem 7 (Orthogonality under Integration)** Let: \((N_{\gamma_1}(t))\) and \((N_{\gamma_2}(t))\) be two independent Poisson processes; \(M_1(t) = N_{\gamma_1}(t) - \gamma_1 t\) and \(M_2(t) = N_{\gamma_2}(t) - \gamma_2 t\) be the corresponding compensated martingales; \((F_1(t))\) and \((F_2(t))\) be almost surely bounded predictable processes with respect to the natural filtrations \((\sigma(M_1(t); t \leq s))_{s \geq 0}\) and \((\sigma(M_2(t); t \leq s))_{s \geq 0}\), respectively. Then,

\[
\left( \int_0^t F_1(s) dM_1(s) \right) \left( \int_0^t F_2(s) dM_2(s) \right)
\]

is a martingale.

**Proof:** Refer to [8].

In particular, it follows as a Corollary to Theorem 7 that

\[
E \left[ \left( \int_0^t F_1(s) dM_1(s) \right) \left( \int_0^t F_2(s) dM_2(s) \right) \right] = E \left[ \left( \int_0^t F_1(s) dM_1(s) \right) \left( \int_0^t F_2(s) dM_2(s) \right) \right] = E \left( \int_0^t F_1(s) dM_1(s) \right) = 0, \ \forall t \geq 0.
\]

In words, the resulting integral martingales \((M_{ai}(t))\) and \((M_{bi}(t))\) are orthogonal. Now, back to the proof of Theorem 6.

\[
E(M_i^N(t))^2 = E \left[ \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t \mathbf{1}_{Y_i^N(s)=\ell,Y_j^N(s)=q} \left( N_{\gamma_1,q} \left( \frac{N_i-\ell}{N_i} \right) ds - \gamma_{ji} q \left( \frac{N_i-\ell}{N_i} \right) ds \right)^2 \right]
= \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} E \left[ \int_0^t \mathbf{1}_{Y_i^N(s)=\ell,Y_j^N(s)=q} \left( N_{\gamma_1,q} \left( \frac{N_i-\ell}{N_i} \right) ds - \gamma_{ji} q \left( \frac{N_i-\ell}{N_i} \right) ds \right)^2 \right]
= \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} E \left[ \int_0^t \mathbf{1}_{Y_i^N(s)=\ell,Y_j^N(s)=q} \gamma_{ji} q \left( \frac{N_i-\ell}{N_i} \right) ds \right]
\leq E \left( \int_0^t \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \mathbf{1}_{Y_i^N(s)=\ell,Y_j^N(s)=q} \gamma_{ji} N_j ds \right)
\leq \gamma_{ji} N_j t,
\]

where the second equality follows from Theorem 7, remarking that all Poisson processes involved, indexed by \(\ell\) and \(q\), are independent and thus the integral martingales are orthogonal and the cross-terms cancel out. The third equality is due to the Itô isometry Theorem (refer to [11] or [12]) and the fact that

May 7, 2014
the quadratic variation of a compensated Poisson martingale is given by \( \langle N_\gamma(t) - \gamma t \rangle = \gamma t \). The last inequality holds since the family of sets \( I_{\ell,q} := \{ s \in \mathbb{R} : Y_j^N(s-) = q, Y_i^N(s-) = \ell \} \), indexed by \( \ell, q \), are disjoint and thus

\[
\sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} 1_{\{Y_j^N(s-) = q, Y_i^N(s-) = \ell \}} = 1 \cup_{\ell,q} \{ Y_j^N(s-) = q, Y_i^N(s-) = \ell \} \leq 1_{[0,1]}(s).
\]

Therefore,

\[
E \left( \frac{\gamma j_i}{N_i} t \right) = \frac{1}{N_i} E(\gamma_j^N(t)) \leq \frac{1}{N_i} \frac{\gamma j_i}{N_j} t.
\]

Thus,

\[
E \left( \frac{\gamma j_i}{N_i} t \right) \rightarrow 0
\]

as \( N_i \to \infty \) and \( N_j \to \infty \), and \( \frac{N_j}{N_i} \to \alpha_{ij} < \infty \). Similarly, the variance of the martingale \( M^N_{i}(t) \) converges to zero. Thus, the martingale vanishes in \( L_2 \) with \( O(1/N_i) \).

Since \( M^N_{i}(t) \) is a martingale, from Doob’s inequality:

\[
P \left( \sup_{0 \leq t \leq T} \left| M^N_{i}(t) \right| > \epsilon \right) \leq \frac{E \left( M^N_{i}(T) \right)^2}{\epsilon^2} \rightarrow 0, \ \forall \epsilon > 0, \ \forall T \geq 0.
\]

That is, \( M^N_{i}(t) \) converges to zero in probability in the space of càdlàg paths with the \( sup \) norm.

The next Theorem is an extension for stochastic processes of the statement that convergence in probability implies convergence in distribution for real valued random variables. The Theorem is Proposition C.5 from [8], and it will imply that \( M^N_{i}(t) \) converges weakly (in the Skorokhod topology) to 0. That is, \( P_{M^N_{i}} \), the probability measure over \( D_{\mathbb{R}}[0,T] \) induced by \( M^N_{i}(t) \), converges weakly to \( \delta_0 \) (Dirac measure about \( m(t) = 0 \)).

**Theorem 8** Let \( Z^N(t) \) be a sequence of càdlàg stochastic processes on the interval \([0,T]\) such that,

\[
P \left( \sup_{0 \leq t \leq T} \left| Z^N(t) - z(t) \right| > \epsilon \right) \rightarrow 0, \ \forall \epsilon > 0,
\]

where \( (z(t)) \) is a deterministic càdlàg function over \([0,T]\). Then,

\[
(Z^N(t)) \Rightarrow (z(t)) \text{ on } [0,T],
\]

i.e., \( Z^N(t) \) converges weakly, for the Skorokhod topology, to \( (z(t)) \). In other words, the sequence of probability measures \( P_{Z^N} \) over \( D_{\mathbb{R}}[0,T] \) induced by \( Z^N(t) \) converges weakly to \( \delta_{(z(t))} \) (Dirac measure about \( (z(t)) \)).
Proof: Refer to proposition C.5 in [8].

Thus, we conclude that the martingale \( \left( \mathcal{M}_N^i(t) \right) \) converges weakly to zero.

D. \( \left( \mathcal{Y}_N^N(t) \right) \) converges weakly

The next Theorem is a stochastic version of Arzelà-Ascoli and will provide sufficient conditions to guarantee the tightness of the sequence \( \left( \mathcal{Y}_N^N(t) \right) \)—just as Arzelà-Ascoli provides sufficient conditions to guarantee tightness of a family of (deterministic) functions.

**Theorem 9** Let \( \left( \mathcal{Y}_N^N(t) \right) \) be a sequence of càdlàg processes. Then, the sequence of probability measures \( P^N \) induced on \( D_R \) by \( \left( \mathcal{Y}_N^N(t) \right) \) is tight and any weak limit point of this sequence is concentrated on the subset of continuous functions \( C_R \subset D_R \) if and only if the following two conditions hold for each \( T > 0 \) and \( \epsilon > 0 \):

\[
\lim_{k \to \infty} \lim_{N \to \infty} \mathbb{P} \left( \sup_{0 \leq t \leq T} \mathcal{Y}_N^N(t) \geq k \right) = 0 \quad \text{(Uniform Boundness)} \quad (20)
\]

\[
\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P} \left( \omega(\mathcal{Y}_N^N, \delta, T) \geq \epsilon \right) = 0 \quad \text{(Equicontinuity)} \quad (21)
\]

where we defined

\[
\omega(x, \delta, T) = \sup \left\{ \sup_{u,v \in [t,t+\delta]} |x(u) - x(v)| : 0 \leq t \leq t + \delta \leq T \right\}.
\]

Proof: Refer to [13].

From Theorems 6 and 8, we have that \( \left( \mathcal{M}_N^N(t) \right) \to 0 \). We are just left to show that our sequence \( \left( \mathcal{Y}_i^N(t) \right) \) meets the requirements in equations \((20)\) and \((21)\) and therefore it is tight (or relatively compact). In other words, it admits a convergent subsequence \( \left( \mathcal{Y}_i^{N*}(t) \right) \to \mathcal{Y}(t) \) where \( \left( \mathcal{Y}(t) \right) \) is almost surely continuous. Indeed, \( \mathbb{P} \left( \sup_{0 \leq t \leq T} \mathcal{Y}_i^N(t) \geq k \right) = 0, \forall k > 1, \) and the first condition holds trivially. The second condition is a stochastic version of the equicontinuity condition in the Arzelà-Ascoli Theorem.

\[
\omega \left( \mathcal{Y}_i^N, \delta, T \right) = \sup \left\{ \sup_{u,v \in [t,t+\delta]} |\mathcal{Y}_i^N(u) - \mathcal{Y}_i^N(v)| : 0 \leq t \leq t + \delta \leq T \right\}
\]

\[
= \sup_{0 \leq t \leq t + \delta \leq T} \left\{ \sup_{u,v \in [t,t+\delta]} \left| \mathcal{M}_i^N(u) - \mathcal{M}_i^N(v) \right| \right\}
\]

\[
+ \int_u^v \gamma_{ji} \frac{N_j}{N_i} \mathcal{Y}_j^N(s-) \left( 1 - \mathcal{Y}_i^N(s-) \right) ds - \int_u^v \mu_i \mathcal{Y}_i^N(s-) ds \right\}
\]

\[
\leq \sup_{0 \leq t \leq t + \delta \leq T} \left\{ \sup_{u,v \in [t,t+\delta]} \left| \mathcal{M}_i^N(u) - \mathcal{M}_i^N(v) \right| \right\} + \gamma_{ji} \frac{N_j \delta}{N_i} \quad (24)
\]

\[
:= \omega_2 \left( \mathcal{Y}_i^N, \delta, T \right).
\]
Although in equation (24) there is an explicit \( j \) dependence, we note that, for a bipartite network, choosing island \( i \) fixes the other island \( j \); because of this, the arguments in \( \omega_2 \left( \overline{Y}_i^N, \delta, T \right) \) in equation (25) do not show explicitly the dependence in \( j \). For any \( \epsilon > 0 \), we have
\[
P \left( \omega \left( \overline{Y}_i^N, \delta, T \right) \geq \epsilon \right) \leq P \left( \omega_2 \left( \overline{Y}_i^N, \delta, T \right) \geq \epsilon \right).
\]
Now, from equation (19) and any \( \alpha > 0 \), we can choose \( N_1 \) and \( N_2 \) large enough so that
\[
P \left( \sup_{0 \leq t \leq T} \left| \overline{M}_i^N(t) \right| > \epsilon \right) \leq \alpha.
\]
and therefore,
\[
P \left( \omega \left( \overline{Y}_i^N, \delta, T \right) \geq \epsilon \right) \leq P \left( \omega_2 \left( \overline{Y}_i^N, \delta, T \right) \geq \epsilon \right) < \alpha
\]
for \( \delta \) small enough. We conclude that \( \left( \overline{Y}_i^N(t) \right) \) is a tight family. In other words, it admits a convergent subsequence \( \left( \overline{Y}_{ij}^N(t) \right) \Rightarrow \left( \overline{Y}_t(t) \right) \) where \( \left( \overline{Y}_t(t) \right) \) is almost surely continuous. In fact, the sequence \( \left( \overline{Y}_i^N(t) \right) \) is not only tight, but it converges to the solution of a deterministic differential equation as shown in the next Theorem 10. The main argument is that any weak accumulation point in the tight sequence \( \left( \overline{Y}_i^N(t) \right) \) should obey the equation
\[
\frac{d}{dt} \overline{Y}_t(t) = F \left( \overline{Y}_t(t) \right)
\]
and from the uniqueness of equation (26), the whole sequence converges.

**Theorem 10** Let \( \left( \overline{Y}_{ij}^N(t) \right) \) be a subsequence converging weakly to \( \left( \overline{Y}_t(t) \right) \) (an almost surely continuous process) with \( \overline{Y}_{ij}^N(0) \Rightarrow \overline{Y}(0) \) and let \( \frac{N_j}{N_i} \rightarrow \alpha_{ij} \in \mathbb{R}^+ \). Then
\[
\overline{Y}_i(t) = \overline{Y}_i(0) + \int_0^t \overline{\sigma}_{ji} \overline{Y}_j(s) \left( 1 - \overline{Y}_i(s) \right) ds - \int_0^t \mu_i \overline{Y}_i(s) ds,
\]
where we defined \( \overline{\sigma}_{ji} := \gamma_{ji} \alpha_{ji} \).

**Proof:** We have the following term by term convergence
\[
\overline{Y}_{ij}^N(t) = \overline{Y}_{ij}^N(0) + \overline{M}_i^N(t) + \int_0^t \gamma_{ji} \overline{Y}_{ij}^N(s) \left( 1 - \overline{Y}_{ij}^N(s) \right) ds - \int_0^t \mu_i \overline{Y}_{ij}^N(s) ds
\]
\[
\downarrow 1 \quad \downarrow 2 \quad \downarrow 3 \quad \downarrow 4 \quad \downarrow 5 \quad .
\]
\[
\overline{Y}_i(t) = \overline{Y}_i(0) + 0 + \int_0^t \overline{\sigma}_{ji} \overline{Y}_j(s) \left( 1 - \overline{Y}_i(s) \right) ds - \int_0^t \mu_i \overline{Y}_i(s) ds
\]
Convergences 1 and 2 hold since we assumed that \( \left( \overline{Y}_{ij}^N(t) \right) \Rightarrow \left( \overline{Y}_i(t) \right) \) and \( \overline{Y}_{ij}^N(0) \Rightarrow \overline{Y}_i(0) \). Convergence 3 holds since \( \overline{M}_i^N(t) \Rightarrow 0 \) as proved before. Moreover, since \( \left( \overline{Y}_t(t) \right) \) is almost surely continuous...
continuous, then \( (\bar{Y}^N_i(t)) \rightarrow (\bar{Y}_i(t)) \) almost surely uniformly over compact intervals. Therefore, given \( T > 0 \),
\[
\int_0^t \gamma_{ji} \frac{N_j}{N_i} \bar{Y}^N_i(s) \left( 1 - \bar{Y}^N_i(s) \right) ds \rightarrow \int_0^t \gamma_{ji} \bar{Y}_j(s) \left( 1 - \bar{Y}_i(s) \right) ds,
\]
almost surely uniformly over the interval \([0, T]\) and convergence 4 follows. The 5th case results similarly. Therefore, for any convergent subsequence with \( (\bar{Y}^{N_p}_i(t)) \Rightarrow (\bar{Y}(t)) \) and \( \bar{Y}^{N_p}(0) \Rightarrow \bar{Y}(0) \), it follows that \( (\bar{Y}(t)) \) is solution of the integral equation (27).

Finally, the next theorem rigorously states the emergent dynamics of the single virus spread over bipartite networks as the number of agents grows large.

**Theorem 11** Let \( (\bar{Y}^N_1(0), \bar{Y}^N_2(0)) \Rightarrow y_0 \in \mathbb{R}^2 \). Then, the normalized sequence \( \left( \bar{Y}^N(t) \right) \) converges weakly to the solution \( (y_1(t, y_0), y_1(t, y_0)) \) of the following ODE:

\[
\frac{d}{dt} y_1(t) = (\tau_{21} y_2(t)) (1 - y_1(t)) - \mu_1 y_1(t) \quad (28)
\]
\[
\frac{d}{dt} y_2(t) = (\tau_{12} y_1(t)) (1 - y_2(t)) - \mu_2 y_2(t) \quad (29)
\]

**Proof:** As the underlying vector field of (27), \( F = (F_1, F_2) : [0, 1]^2 \rightarrow \mathbb{R}^2 \), where \( F_i(y_1, y_2) := \gamma_{ji} y_j (1 - y_i) - \mu_i y_i \), is Lipschitz, the continuous (and thus, differentiable) solution \( (\bar{Y}(t)) \) of (27) is unique. Thus, any weak limit of \( \left( \bar{Y}^N(t) \right) \) with initial condition given by \( (\bar{Y}^N(0)) \) and converging weakly to \( \bar{Y}(0) \) is equal to the unique solution \( (\bar{Y}(t)) \) of (27) with initial condition \( (\bar{Y}(0)) \). Therefore, the whole sequence converges \( (\bar{Y}^N(t)) \Rightarrow (\bar{Y}(t)) \) to the solution of (27). Equation (27) is the integral version the ODE (28)-(29). Theorem (11) is concluded.

We showed in this Section that the sequence \( (\bar{Y}^N(t)) \) over the corresponding sequence of bipartite networks \( G^{(N_1, N_2)} \) converges weakly to the solution of a vector differential equation. We explored the martingale structure of the perturbing noise on the dynamics of the process \( (\bar{Y}^N(t)) \) to show that it converges weakly under the Skorokhod topology to zero. As a Corollary to this fact, the family \( (\bar{Y}^N(t)) \) is tight with a single accumulation point given by the (unique) solution of a limiting differential equation. Since any convergent subsequence converges to the same accumulation point, then the whole sequence \( (\bar{Y}^N(t)) \) converges to the unique accumulation point, namely, the solution of the ODE (28)-(29). In the next section, we extend the convergence result to the multi-virus multipartite network case.

**IV. Mean Field – Multivirus over Multipartite Networks**

In Section III, we established that the sequence of single virus macroprocesses \( \left( \bar{Y}^N_1(t), \bar{Y}^N_2(t) \right) \) over the corresponding sequence of bipartite networks \( G^N \) converges weakly to the solution \( (y_1(t), y_2(t)) \) of
a deterministic ODE given by equations (28)-(29). We divided the proof into four main steps:

(i) The martingale \((\overline{M}_1^N(t), \overline{M}_2^N(t))\) converges weakly to zero;
(ii) The family \((\overline{Y}_1^N(t), \overline{Y}_2^N(t))\), indexed by \(N\), is tight;
(iii) Any accumulation point of the tight family is solution of (28)-(29);
(iv) Uniqueness of the differential equation (28)-(29) implies convergence.

In this Section, we extend Theorem 11 to the more general case of multi-virus epidemics over multipartite networks. We consider in Subsection IV-A single-virus over multipartite networks and then the general case of multivirus over multipartite networks is in Subsection IV-B.

A. Single-virus over Multipartite Networks

For single virus spread, remark that \(Y^N(t) = (Y_1^N(t), \ldots, Y_M^N(t))\) stands for the process associated with the fraction of infected nodes at each island \(i \in \{1, \ldots, M\}\) over the multipartite network \(G^N\) with \(M\) islands and supertopology induced by the topology of the graph \(G\). In Section III, we obtain the following pathwise description for \(Y^N(t)\)

\[
Y_i^N(t) = Y_i^N(0) + M_i^N(t) + \sum_{j \sim i} \int_0^t \gamma_{ji} Y_j^N(s-) \left( \frac{N_i - Y_i^N(s-)}{N_i} \right) ds - \int_0^t \mu_i Y_i^N(s-) ds,
\]

for \(i = 1, \ldots, M\). The infections from all neighboring islands are now coupled by these \(M\) equations.

The corresponding martingales are given by

\[
M_i^N(t) = \sum_j \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} \int_0^t 1_{\{Y_i^N(s-)=q, Y_j^N(s-)=\ell\}} \left( N_{\gamma_{ji}q(\frac{N_i-\ell}{N_i})} ds - \gamma_{ji} q \left( \frac{N_i - \ell}{N_i} \right) ds \right)
.
\]

Next, we prove that the sequence \((Y^N(t))\) over the underlying sequence of multipartite networks \(G^N\) converges weakly to the solution of an ODE.

**Theorem 12** Let \((Y^N(0)) \Rightarrow y_0 \in \mathbb{R}^M\) with \(N = (N_1, \ldots, N_M) \to \infty\) and \(\frac{N_i}{N_j} \to \alpha_{ij} < \infty\) for all \(i \sim j\). Then, the normalized sequence \((\overline{Y}^N(t))\) converges weakly to the solution \((y(t,y_0)) = \ldots\)
(y_1(t, y_0), \ldots, y_M(t, y_0)) of the following ODE:

\[
\frac{d}{dt} y_i(t) = \left( \sum_{j \sim i} \gamma_{ji} y_j(t) \right) (1 - y_i(t)) - \mu_i y_i(t) := F_i(y_1(t), \ldots, y_M(t))
\]

with \( \gamma_{ji} := \alpha_{ji} \gamma_{ji} \).

**Proof:** For the sake of clarity, we revisit each of the points (i)-(iv) referred to in the beginning of this Section.

(i) **Martingale vanishes.** We start by observing that, for fixed \( i, j \in \{1, \ldots, M\} \), the underlying Poisson point processes \( N_{\gamma,jq}(\frac{N_i - t}{N_i}) \) indexed by \((\ell, q) \in \{1, \ldots, N_i\} \times \{1, \ldots, N_j\}\) are independent as explained in Section III-B. Now, fixing only \( i \), the Poisson processes are still independent due to the independence of the exponential time services associated with different nodes. Therefore, it follows as a Corollary to Theorem 7 that, for fixed \( i \), the martingales \( M_{aij}^N(t, \ell, q) \), indexed by \( j, \ell, q \) are orthogonal in the same sense as in equations (17)-(18). Let \( M_{ai}^N(t) := \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} M_{aij}^N(t, \ell, q) \). It turns out that

\[
E \left( M_{ai}^N(t) \right)^2 = E \left( \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} M_{aij}^N(t, \ell, q) \right)^2 = \sum_{j \sim i} \sum_{\ell=1}^{N_i} \sum_{q=1}^{N_j} E \left( M_{aij}^N(t, \ell, q) \right)^2 \leq \sum_{j \sim i} \gamma_{ji} N_j t \leq M \max_{j=1, \ldots, M} \{ \gamma_{ji} N_j \} t,
\]

and thus,

\[
E \left( M_{ai}^N(t) \right)^2 = \frac{1}{N_i^2} E \left( M_{ai}^N(t) \right)^2 \leq \frac{M}{N_i} \left( \max_{j=1, \ldots, M} \left\{ \frac{\gamma_{ji} N_j}{N_i} \right\} \right) t \to 0.
\]

Similarly \( E \left( M_{bi}^N(t) \right)^2 \to 0 \) and therefore, the normalized martingale \( M_{ai}^N(t) \) converges to zero in \( L^2 \) for all time \( t, t \geq 0 \). Now, from Doob’s inequality and from Theorem 8, we conclude that \( \left( M_{ai}^N(t) \right) \to 0. \)
(ii) The family \( \{ \bar{Y}^N(t) \} \) is tight. As in equations (22)-(25), we have

\[
\omega \left( \bar{Y}^N_i, \delta, T \right) \leq \sup_{0 \leq t \leq t + \delta \leq T} \left\{ \sup_{u, v \in [t, t + \delta]} \left| M_i^N(u) - M_i^N(v) \right| \right\} + \sum_{j \sim i} \gamma_{ji}^N \frac{N_i}{N_j} \delta \quad (31)
\]

\[
:= \omega_2 \left( \bar{Y}^N_i, \delta, T \right). \quad (32)
\]

From Theorem 9 and similar arguments as in Section III, we conclude that \( \{ \bar{Y}^N(t) \} \) is a tight family, that is, it admits a convergent subsequence \( \{ \bar{Y}^{N_k}(t) \} \Rightarrow (Y(t)) \). Also, from Theorem 9, \( (Y(t)) \) is almost surely continuous.

(iii) If \( \{ \bar{Y}^{N_k}(t) \} \Rightarrow (Y(t)) \) then, \( (Y(t)) \) is solution of the ODE (30). It follows similarly to as done in the proof of Theorem (10), remarking that we assume a finite (fixed) number of islands \( M \).

(iv) \( \{ \bar{Y}^N(t) \} \Rightarrow (Y(t)) \), where \( (Y(t)) \) is solution of the ODE (30). Note that the underlying vector field

\[
F \left( y_1, \ldots, y_M \right) = (F_1 \left( y_1, \ldots, y_M \right), \ldots, F_M \left( y_1, \ldots, y_M \right)) \quad (33)
\]

in equation (30) is differentiable and thus, \( F \) is locally Lipschitz. Therefore, solution of (30) exists locally and it is unique. Since the state space of interest \([0, 1]^M\) is compact and invariant, \( F \) is globally Lipschitz over \([0, 1]^M\), and any solution of (30) is defined for all time \( t \geq 0 \) and is unique. In particular, any convergent subsequence converges to the same weak limit given by the unique solution of (30) and, thus, the whole sequence converges. This concludes the proof of Theorem 12.

B. Multivirus over Multipartite Networks

We denote as \( \{ \bar{Y}^N(t) \} = \left[ \bar{Y}^N_{ik}(t) \right]_{ik} \) the matrix process collecting the fraction of \( k \)-infected nodes at island \( i \in \{1, \ldots, M\} \) with \( k \in \{1, \ldots, K\} \) over time \( t, t \geq 0 \), where \( K \) is the number of virus strains. In this Subsection, we refer to \( \{ \bar{Y}^N_i(t) \} = \left( \bar{Y}^N_{i1}(t), \ldots, \bar{Y}^N_{iK}(t) \right) \) as the distribution of infected nodes at island \( i \) across the \( K \) strains of virus. Recall the definition of the state-space \( \mathcal{Y}^N_i \) of the process \( \left( \bar{Y}^N_i(t) \right) \) given in Section III-A. Applying the same reasoning as in Section III-B, we obtain the following pathwise description for \( \left( \bar{Y}^N_i(t) \right) \)

\[
Y_{ik}^N(t) = Y_{ik}^N(0) + M_{ik}^N(t) + \sum_{j \sim i} \int_0^t \gamma_{ji}^k Y_{jk}^N(s-) \left( \frac{N_i - \langle Y_i^N(s-), 1 \rangle}{N_i} \right) ds - \int_0^t \mu_i^k Y_{ik}^N(s-) ds,
\]
where the martingale is given by

\[
M_{ik}^N(t) = \sum_{j \sim i} \sum_{y \in \mathcal{Y}_N} \sum_{q=1}^{N_j} \int_0^t 1\{Y^N_{jk}(s-)=q, Y^N_{ki}(s-)=y\} \left( \mathcal{N}_{\gamma_{ji} q} \left( \frac{N_i - (y, 1)}{N_i} \right) (ds) - \gamma_{ji} q \left( \frac{N_i - (y, 1)}{N_i} \right) ds \right) \]

\[
- \sum_{\ell=1}^{N_i} \int_0^t 1\{Y^N_{ki}(s-)=\ell\} (\mathcal{N}_{\ell \mu_i} (ds) - \ell \mu_i ds)
\]

and from the construction in Section III-B, \( \{\mathcal{N}_{\gamma_{ji} q} \left( \frac{N_i - (y, 1)}{N_i} \right) \}_{i,j,q,y} \) is a family of independent Poisson processes.

**Theorem 13** Let \( (\mathbf{Y}^N(0)) \Rightarrow \mathbf{y}_0 \in \mathbb{R}^{M \times K} \) with \( N = (N_1, \ldots, N_M) \to \infty \) and \( \frac{N_i}{N_M} \to \alpha_{ij} < \infty \) for all \( i \sim j \). Then, the normalized sequence \( (\mathbf{Y}^N(t)) \) converges weakly to the solution \( (\mathbf{y}(t, \mathbf{y}_0)) \) of the following ODE:

\[
\frac{d}{dt} y_{ik}(t) = \left( \sum_{j \sim i} \gamma_{ji}^k b_{jk}(t) \right) \left( 1 - \sum_{m=1}^{K} y_{im}(t) \right) - \mu_i^k y_{ik}(t)
\]

with \( \gamma_{ji}^k := \alpha_{ji} \gamma_{ji}^k \).

**Proof:**

i) **Martingale vanishes.** Since the underlying Poisson point processes \( \mathcal{N}_{\gamma_{ji} q} \left( \frac{N_i - (y, 1)}{N_i} \right) \) are independent then, it follows as a Corollary to Theorem 7 that the compensated martingales \( M_{j,y,q}^N(t) \), are pairwise orthogonal. It turns out that

\[
E \left( \sum_{j \sim i} \sum_{y \in \mathcal{Y}_N} \sum_{q=1}^{N_j} M_{j,y,q}^N(t) \right)^2 = \sum_{j \sim i} \sum_{y \in \mathcal{Y}_N} \sum_{q=1}^{N_j} E \left( M_{j,y,q}^N(t) \right)^2
\]

\[
= \sum_{j \sim i} \sum_{y \in \mathcal{Y}_N} \sum_{q=1}^{N_j} \left( \int_0^t 1\{Y^N_{jk}(s-)=q, Y^N_{ki}(s-)=y\} \left( \gamma_{ji}^k \left( \frac{N_i - (y, 1)}{N_i} \right) ds \right) \right)
\]

\[
\leq \sum_{j \sim i} \left( \int_0^t \sum_{y \in \mathcal{Y}_N} \sum_{q=1}^{N_j} 1\{Y^N_{jk}(s-)=q, Y^N_{ki}(s-)=y\} \gamma_{ji}^k N_j ds \right)
\]

\[
\leq M_{j=1,\ldots,M} \left\{ \gamma_{ji} N_j \right\} t,
\]
and thus,
\[
\mathbb{E} \left( \frac{1}{N_i} \sum_{j \sim i} \sum_{y \in Y_i^N} \sum_{q=1}^{N_j} M_{j,y,q}^N(t) \right)^2 \leq \frac{M}{N_i} \left( \max_{j=1,\ldots,M} \left\{ \frac{\gamma_{ji}^k N_j}{N_i} \right\} \right) t \to 0.
\]
Similarly \( \mathbb{E} \left( \overline{M}_b^N(t) \right)^2 \to 0 \) and therefore, the normalized martingale \( \overline{M}^N_{ik}(t) \) converges to zero in \( \mathcal{L}_2 \) for all time \( t \), \( t \geq 0 \). Now, from Doob’s inequality and from Theorem 8, we conclude that \( \left( \overline{M}^N_{ik}(t) \right) \Rightarrow 0 \).

ii) The family \( \overline{Y}^N(t) \) is tight. As in equations (22)-(25), we have
\[
\omega \left( \overline{Y}^N_{ik}, T, \delta \right) \leq \sup_{0 \leq t \leq t+\delta \leq T} \left\{ \sup_{u,v \in [t,t+\delta]} \left| M_{ik}^N(u) - M_{ik}^N(v) \right| \right\} + M \frac{\max_{j=1,\ldots,M} \left\{ \gamma_{ji}^k N_j \right\}}{N_i} \delta.
\]
From Theorem 9 and similar arguments as in Section III, we conclude that \( \left( \overline{Y}^N(t) \right) \) is a tight family, that is, it admits a convergent subsequence \( \left( \overline{Y}^N_{ik}(t) \right) \Rightarrow (\overline{Y}(t)) \). Also, from Theorem 9, \( (\overline{Y}(t)) \) is almost surely continuous.

iii) If \( \left( \overline{Y}^N_{ik}(t) \right) \Rightarrow (\overline{Y}(t)) \) then, \( (\overline{Y}(t)) \) is solution of the ODE (34). It follows similarly to as done in the proof of Theorem (10), remarking that we assume a finite (fixed) number of islands \( M \).

iv) \( \left( \overline{Y}^N(t) \right) \Rightarrow (\overline{Y}(t)) \), where \( (\overline{Y}(t)) \) is solution of the ODE (34). Note that for similar reasons as exposed in Subsection IV-A, solution of (34) exists and is unique. Thus, any convergent subsequence \( \left( \overline{Y}^N_{ik}(t) \right) \) converges to the same weak limit given by the unique solution of (34) and thus, the whole sequence converges.

V. Conclusion

In this paper, we established the fluid limit dynamics of a multivirus epidemics over a multipartite network from a peer-to-peer stochastic network model of diffusion. Namely, we proved that the normalized macrostate \( \left( \overline{Y}^N_{ij}(t) \right) \) collecting the fraction of \( j \)-infected nodes \( \left( \overline{Y}^N_{ij}(t) \right) \) per island \( i \in \{1, \ldots, M\} \) with \( j \in \{1, \ldots, K\} \) over \( G^N \) converges weakly, under the Skorokhod topology on the space of càdlàg sample paths, to the solution \( (y(t)) \) of a \( (M \times K) \)-dimensional ordinary differential equation given by (34). To this effect, we first proved that the underlying martingale perturbation \( \left( \overline{M}^N(t) \right) \) vanishes as \( N \) grows large, which implies that the macrostate family \( \left( \overline{Y}^N_{ij}(t) \right) \) is tight in \( N \). Then, we showed that any weak accumulation point of the family \( \left( \overline{Y}^N(t) \right) \) is solution to the vector ordinary differential equation (34) with Lipschitz vector field. From the uniqueness of the solutions of the resulting meanfield differential equation (34), we concluded that the whole sequence \( \left( \overline{Y}^N_{ij}(t) \right) \) converges weakly to the solution of (34). We now present a numerical experiment of two strains of virus \( x \) and \( y \) spreading across a bipartite network via our SIS stochastic peer-to-peer law of infection. Figure 4 illustrates the Matlab results for
the evolution of two strains in the bipartite network (refer to the noisy curves) and we superimpose on it the corresponding meanfield evolution (refer to the smooth curves). Figures 4a, 4b, and 4c illustrate the evolution of the fractions of $x$-infected (blue/solid curves) and $y$-infected (red/dashed curves) nodes at islands 1 and 2. The boldfaced curves represent the associated meanfield solutions. We observe that,

![Figure 4: The plots represent the numerical evolution of the fractions of $x$-infected (in blue or solid) and $y$-infected (in red or dashed) nodes at each island 1 and 2. The boldfaced curves represent the solution of the limiting vector differential equation of a bi-viral epidemics in a bipartite network.](image)

as the number of nodes grows large, the randomness decreases and the infected population dynamics fits the meanfield prediction. For the experiment, we have set $\gamma^x := \gamma^x_{12} = \gamma^x_{21} > \gamma^y_{12} = \gamma^y_{21} =: \gamma^y$ with $\mu^x = \mu^y = 1$ and observe, in particular, that the most virulent strain survives and the weaker strain dies out. The qualitative analysis of the meanfield dynamics (34) is developed in [14] where we proved that the natural selection phenomenon observed in Figures 4a, 4b, and 4c extends to symmetric regular multipartite networks, where by symmetric and regular we mean $\gamma^y_{ij} = \gamma^y_{ji}$ and $d = d_i = d_j$ for all pair of communicating islands $i$ and $j$ and for all types of virus $y$.

**REFERENCES**

[1] E. Olivieri and M. E. Vares, *Large Deviations and Metastability*, ser. Encyclopedia of Mathematics and its Applications. London, UK: Cambridge University Press, 2005, no. 100.

[2] F. Cucker and S. Smale, “Emergent behavior in flocks,” *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 852–862, May 2007.

[3] S. Kar and J. M. F. Moura, “Global emergent behaviors in clouds of agents,” in *2011 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011, pp. 5796–5799.

[4] M. P. Richardson, “Large scale brain models of epilepsy: dynamics meets connectomics,” *Journal of Neurology, Neurosurgery, and Psychiatry*, no. 83, pp. 1238–1248, December 2012.

[5] N. Antunes, C. Fricker, P. Robert, and D. Tibbi, “Stochastic networks with multiple stable points,” *The Annals of Applied Probability*, vol. 36, no. 1, pp. 255–278, 2008.

[6] P. Van Mieghem, J. Omic, and R. Kooij, “Virus spread in networks,” *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, Feb. 2009.
[7] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales: Foundations*, 2nd ed., ser. Cambridge Mathematical Library. Cambridge, UK: Cambridge University Press, April 2000, vol. 1.

[8] P. Robert, *Stochastic Networks and Queues*, ser. Stochastic Modelling and Applied Probability. Springer-Verlag, 2003, vol. 52.

[9] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Probability and Statistics. John Wiley & Sons, Inc., 1995.

[10] A. Santos and J. M. F. Moura, “Emergent behavior in large scale networks,” in *2011 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, December 2011, pp. 4485 –4490.

[11] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales: Foundations*, 2nd ed., ser. Cambridge Mathematical Library. Cambridge, UK: Cambridge University Press, September 2000, vol. 2.

[12] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 1st ed., ser. Graduate Texts in Mathematics. New York: Springer-Verlag, September 1988, vol. 113.

[13] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., ser. Grundlehren der Mathematischen Wissenschaften. New York, NY: Springer-Verlag, 2003, vol. 288.

[14] A. Santos, J. M. F. Moura, and J. M. F. Xavier, “Epidemics in multipartite networks: Emergent dynamics,” July 2013, submitted.