2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Emergent Behavior in Large Scale Networks

Augusto Santos José M. F. Moura
Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh PA 15213-3890

Abstract— **We apply mean field asymptotic analysis to explain the emergence of global behavior in large scale networks. The underlying motivating application is epidemics like computer virus spreading, for example, in wide campus local networks. We consider multiple classes of viruses, each type bearing their own statistical characterization – exogenous contamination, contagious propagation, and healing. The network state (distribution of nodes infected by each class in the network) is a jump Markov process, not necessarily reversible, making it a challenge to obtain its invariant distribution. By suitable renormalization, in the limit of a large network (number of nodes,) the macroscopic behavior of the network is described by the solution of a set of deterministic nonlinear differential equations (Riccati type.) We show that, under the heavy traffic assumption, the relevant underlying dynamics induces a coherent nontrivial metastable behavior in a macroscopic space-time scale: a slight imbalance on the effective spreading rate of one class over the others determines a significantly greater steady state predominance of this class over the others, regardless of the initial distribution.**

I. Introduction

This paper considers the emergence of global behaviors in large scale networks from the microscopic interactions of the networked agents. The global behavior corresponds to the asymptotic state of their network derived by a mean field analysis in the limit of a large number of agents. This behavior is described as the solution of a nonlinear ordinary vector differential equation of the logistic type (or of the Riccati type.) The equilibria of this nonlinear equation is highly sensitive to values of the parameters that govern the interactions among agents. To be concrete, we consider applications in epidemics of multiple types of viruses infecting a closely knitted population, for example, multiple classes of virus contaminating the computers and servers in a campus local area network. We describe the evolution of the fractions of infected nodes of each class on the network. We assume that the viruses enter the network according to a Poisson process with rate λ_i (indexed by the type of virus.) The epidemic then spreads via contamination among the nodes. The nodes may heal after some random time. We assume exponential distributions for the contamination and healing times with rates γ_i and μ_i, respectively. Within this model, our state process is described by a Markov jump process (see [10].) We analyze its asymptotic behavior when the number of nodes N goes to infinity. Namely, we prove in section IV-C that, in the limit of N large, the Markov process converges to the solution of a (deterministic) ordinary vector differential equation $\frac{d}{dt}\mathbf{y}(t) = \mathbf{F}(\mathbf{y}(t))$ for $t \geq 0$, where the vector field $\mathbf{F}(\cdot)$ depends on the dynamics of the viruses. We show through numerical simulations that small perturbations on the statistics of the different types of viruses may lead to completely different steady-state configurations, namely, a small advantage on the effective transmission rate $\tau_i = \frac{\gamma_i}{\mu_i}$ of a particular virus may be key for its class to go from a state of almost extinction to a possibly overwhelming dominance of the network. In section IV, we detail our model and construct the Markov jump process by characterizing its rate transition matrix $\mathbf{Q}^N$. We write down the underlying stochastic integral equation and prove its mean field property. In section V, we study the qualitative behavior of $(\mathbf{y}(t))$ to show metastability, and in section VI, we present numerical simulations that illustrate our analysis. In section VII, we conclude the paper.

II. Brief Review of the Literature

Our underlying application is the diffusion of viruses in networks, though the resulting model may describe epidemiological phenomena in general, as well as diffusion of information, e-mail spam, spread of gossip, etc. Our model resembles the susceptible-infected-susceptible (SIS) model (refer to Bailey [6]) where a node must be in one of two possible states: infected or not infected, but susceptible to infections. Beyond SIS, there are external infections, and we allow for different types of infections (details in section IV.) Another standard model is the SIR model (see [8]) whose main difference from SIS is that a node either heals from an infection and becomes immune and non-susceptible or it dies. A reference discussing the above applications and diffusion models is, for example, [7]. Reference [4] studies the influence of network topology on the spread of a single class of virus. It proposes the N-intertwined mean field model to tie the dynamics of the probabilities of infection

This work was partially supported by NSF grant #CCF1011903 and AFOS grant #A95501010291.

Augusto Santos support was provided by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the Carnegie Mellon Portugal Program under Grant SFRH / BD / 33516 / 2008.

Augusto Santos is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, USA, and Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal. augustos@andrew.cmu.edu

José M. F. Moura is with the Department of Electrical and Computer Engineering, Carnegie Mellon University, USA. moura@ece.cmu.edu, ph: (412)268-6341

of each node to topology features. It shows an epidemic threshold $\tau_c = 1/\lambda_{\max}(\mathbf{A})$ where $\lambda_{\max}(\mathbf{A})$ represents the maximum eigenvalue associated to the adjacency matrix $\mathbf{A}$ of the network. When $\frac{\gamma}{\mu} > \tau_c$ the virus is perpetuated in the network, whereas for $\frac{\gamma}{\mu} < \tau_c$ it exponentially dies out, where γ and μ are the rate of contamination (per link in [4]) and healing, respectively. Our own studies, in the special case of no external arrivals extend and modify the result for multiple virus infection. Letting γ_i and μ_i be the contamination and healing rates, respectively, for virus class i, we show that if $\tau_i = \frac{\gamma_i}{\mu_i} < 1$ then virus type i will die out and if $\tau_i = \frac{\gamma_i}{\mu_i} > 1$ then it survives only if $\tau_i \geq \tau_j$ for all $j \neq i$. A main difference between ours and these previous approaches is that we assume multiple classes of virus spreading homogeneously across the network (an infected node contaminates any other healthy one chosen uniformly at random.) Our work is related to reference [1] that also considers multiclass flow of packets over a complete network with finite capacity sites. The state Markov vector process $\left(\overline{\mathbf{Y}}^N(t)\right)$ (fraction of packets for the different classes in the network) does not model a diffusion process but rather walks of finite size packets. Reference [1] proves that in the limit (as the number of nodes N goes to infinity) $\left(\overline{\mathbf{Y}}^N(t)\right)$ converges to the solution of a deterministic differential vector equation that admits, under appropriate conditions on the size and statistical dynamics of the packets, two stable equilibrium points and a saddle point. This reference calls this configuration as metastable in the sense that the system stays a long period of time in one of the stable equilibrium states and then, due to (rare) perturbations on $\left(\overline{\mathbf{Y}}^N(t)\right)$, the system drifts via the saddle point towards the stronger equilibrium. Our concept of metastability differs from [1] (or [4]). Metastability in this paper stands for sharp changes on the equilibrium state due to perturbations on the statistical parameters rather than on the process $\left(\overline{\mathbf{Y}}^N(t)\right)$ itself. Reference [3] considers the problem of multiclass diffusion in a sparse connected network of densely connected supernodes. Reference [13] presents an overview over contact networks epidemiology.

III. Notation

We summarize in this section the relevant notation used throughout this paper. For the most part, we adopt the notation in [1]. Reference [11] provides background on Markov and diffusion processes.

- $\mathbf{1}_m \in \mathbb{R}^m$: vector with all entries equal to one. We will omit the subscript m whenever clear from the context.
- $\mathbb{N} = \{0,1,2,\ldots\}$: set of natural numbers.
- $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$: set of integers.
- $\mathbb{R}^n_{++}$: positive orthant.
- $\overset{\circ}{A} \subset \mathbb{R}^n$: interior of set A.
- $\mathbf{e}_i$: canonical vector with all entries equal to zero but the i-th entry.
- $\mathbf{1}_{\{A\}}$: the characteristic function for set A. $\mathbf{1}_{\{A\}}(x) = \mathbf{1}_{\{x \in A\}} = 1$ if $x \in A$ and 0 otherwise.
- $T \sim \mathsf{Exp}(\upsilon)$: T is a random variable exponentially distributed with parameter υ.
- ***i*-infected node**: node contaminated with virus type i.
- $\left(Y_i^N(t)\right)$: the Markov jump process accounting for the number of i-infected nodes in a complete network with N nodes. Also, $\left(Y_i^N(t)\right)_{a \leq t \leq b}$ represents the restriction of $\left(Y_i^N(t)\right)$ to the interval $[a,b]$.
- $Y_i^N(t)$: number of i-infected nodes in a complete network with N nodes at time t.
- $\mathbf{Y}^N(t) = \left[Y_1^N(t)\, Y_2^N(t) \ldots Y_K^N(t)\right]^\top \in \mathbb{R}^K$: vector stacking the number of infected nodes for the different classes. $\left(\mathbf{Y}^N(t)\right)$ represents its associated stochastic process.
- $\chi = \left\{(n_1,\ldots,n_K) \in \mathbb{N}^K : \sum_{j=1}^K n_j \leq N\right\}$: phase or state space of $\left(\mathbf{Y}^N(t)\right)$. The state $(n_1,n_2,\ldots,n_K)$ represents the number of nodes n_1 up to n_K infected by each virus from class 1 to class K.
- $\left(\overline{Y}_i^N(t)\right)$: the Markov jump process associated to the fraction of i-infected nodes in a complete network with N nodes at time t. It is defined as $\overline{Y}_i^N(t) = \frac{Y_i^N(t)}{N}$.
- $\Delta = \left\{\mathbf{y} \in \mathbb{R}^K : \mathbf{y}^\top \mathbf{1} \leq 1, \mathbf{y} \geq 0\right\}$: simplex defining the phase space of $\left(\overline{\mathbf{Y}}^N(t)\right)$. Each point $\mathbf{y} \in \Delta$ represents the empirical distribution of the fraction of infected nodes over the different classes of virus.
- $\mathbf{Q}^N$: rate transition matrix associated to the Markov process $\left(\mathbf{Y}^N(t)\right)$. It is defined as:

$$\mathbf{Q}^N(\mathbf{u},\mathbf{v}) = \lim_{t \to 0} \frac{P\left(\mathbf{Y}^N(t) = \mathbf{v} | \mathbf{Y}^N(0) = \mathbf{u}\right)}{t}$$

for $\mathbf{u}$, $\mathbf{v} \in \chi$ and $\mathbf{u} \neq \mathbf{v}$. And

$$\mathbf{Q}^N(\mathbf{u},\mathbf{u}) = -\sum_{\mathbf{v} \in \chi, \mathbf{v} \neq \mathbf{u}} \mathbf{Q}^N(\mathbf{u},\mathbf{v}).$$

- $(\Omega, \mathscr{F}, P)$: probability space underlying the arrival, contamination and healing processes: Ω is the sample space, $\mathscr{F}$ the σ-algebra, and P is the probability measure.

IV. Stochastic Model

A stream of arrivals of K different classes of viruses launches an epidemic over a closely knitted population of N individuals. For the sake of simplicity, we define the individuals as nodes (or sites) in a complete network. Each infected node can transmit the virus to any of the non-infected remaining $N-1$ ones. We assume that a node can either be healthy or infected by only a single type of virus. Although simplistic, the model is rich enough to illustrate the applicability of our method and some of the relevant phenomena that arise. We summarize the relevant processes and parameters.

- **Arrival in the network:** The process that counts the number of type i *virions*[1] entering a node is a Poisson point process with rate λ_i. This is equivalent to assuming that virus type i arrives at the network with rate $\lambda_i N$ and lands uniformly randomly at any node. If the node is already infected, the new arrival has no effect.
- **Healing:** An i-infected node takes time $T \sim \mathsf{Exp}(\mu_i)$ to heal.

[1] Virus particle.

• **Contamination:** Once a node is i-infected, it takes time $T \sim \mathsf{Exp}(\gamma_i)$ to contaminate one (randomly chosen) of the remaining nodes in the network. There is no effect if the chosen node is already infected.

• **Single infection:** If a node is infected, it cannot be re-infected by any other virus, before it heals.

We start by characterizing the process $\left(Y_i^N(t)\right)$ that represents the number of i-infected nodes in a network with N nodes over time $t \geq 0$, for $i \geq 1$. We define $(Y_0^N(t))$ as the process accounting for the number of healthy nodes, i.e., $Y_0^N(t) = N - \sum_{i=1}^K Y_i^N(t)$ for all $t \geq 0$. Due to the memoryless nature of the stochastic processes involved (arrivals, contamination, and healing,) $\left(Y_i^N(t)\right)$ is a Markov process.

A. Dynamics of Stochastic Systems

For a deterministic dynamical system we recall:

• **Phase space:** Set of all reachable states $\mathscr{S}$.

• **Evolution law:** Family of maps $\phi_t : \mathscr{S} \to \mathscr{S}$, where $\phi_t(s_0)$ represents the state of the system after time $t \geq 0$ provided it started at $s_0 \in \mathscr{S}$. In general $(\phi_t(s_0))_{t\geq 0}$ is the solution of a differential equation over $\mathscr{S}$.

For stochastic systems, the evolution law $\Phi_t : \mathscr{P}(\mathscr{S}) \to \mathscr{P}(\mathscr{S})$ is stochastic, where $\mathscr{P}(\mathscr{S})$ is the set of probability measures over (a σ-algebra defined on) $\mathscr{S}$. Now, $\Phi_t(\mu_0)$ represents the probability distribution of the state of the system after time t provided the initial distribution is $\mu_0 \in \mathscr{P}(\mathscr{S})$. In general, $(\Phi_t(\mu_0))_{t\geq 0}$ is the solution of a differential equation over $\mathscr{P}(\mathscr{S})$.

In our problem, the system is stochastic and the (finite) phase space is $\mathscr{S} = \chi$. Its state is given by the stochastic process $\left(\mathbf{Y}^N(t)\right)$. Characterizing its dynamics means specifying the differential equation governing its probability distribution over time. For a general Markov process $(\mathbf{Y}(t))$ on a finite phase space $\chi = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$, it can be shown that the stochastic evolution law is given by $\Phi_t(\mathbf{p}_0) = \mathbf{p}_0^\top e^{\mathbf{Q}t}$, where the matrix $\mathbf{Q} = (q_{ij}) \in \mathbb{R}^{K\times K}$ is the rate transition matrix defined in section III. $\mathbf{Q}$ summarizes the dynamics of the Markov process of interest. For details, refer to chapter 3 of [11]. Next, we compute the rate transition matrix $\mathbf{Q}$ and derive the integral equation satisfied by $\left(\mathbf{Y}^N(t)\right)$. This integral equation is then applied to prove the mean field property.

B. $\mathbf{Q}^N$-Matrix Characterization

We now characterize the stochastic evolution law of the Markov process $\left(\mathbf{Y}^N(t)\right)$ by specifying its rate transition matrix $\mathbf{Q}$. Since the arrivals, transmissions, and healing induce independent Poisson simple processes, jumps greater than one happen with probability zero (refer to [10].) In words, two or more events cannot happen at the same time, where by event we mean either a node healing or becoming infected. When $Y_i^N(t) = u_i$, that is, there are u_i i-infected nodes present in the network, the first healing time T^h is the minimum of u_i independent exponentially distributed random variables with parameter μ_i

$$T^h = \min\{T_1, T_2, \ldots T_{u_i}\} \sim \mathsf{Exp}(\underbrace{\mu_i + \mu_i + \ldots + \mu_i}_{u_i \text{ times}}).$$

Therefore, $\mathbf{Q}^N(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - \mathbf{e}_i) = \mu_i \mathbf{Y}_i^N(t) \mathbf{1}_{\{\mathbf{Y}^N(t)-\mathbf{e}_i \in \chi\}}$. Each infected node transmits a virus to a healthy one with probability $p(t) = Y_0^N(t)/(N-1)$. Remark that a *virion* is transmitted to a random neighbor after a time $T \sim \mathsf{Exp}(\gamma_i)$, but if a neighbor is already infected (which happens with probability $q(t) = 1 - p(t)$,) then the state remains the same. Therefore, with rate $\gamma_i p(t)$ the population of i-infected nodes will be incremented by one due to the presence of 1 infected node. Since we have at time t, $Y_i^N(t)$ i-infected nodes, the final rate due to transmission is given by $\gamma_i Y_i^N(t) p(t)$. The arrival rate is justified in a similar way to be $\lambda_i N \widetilde{p}(t) = \lambda_i (N - \mathbf{1}^\top \mathbf{Y}^N(t))$, where $\widetilde{p}(t)$ is the probability that a virion entering the network lands at a healthy node. In summary, for $\mathbf{Y}^N(t) \in \chi$, we have

$$\begin{cases} \mathbf{Q}^N(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - \mathbf{e}_i) = \mu_i Y_i^N(t) \mathbf{1}_{\{\mathbf{Y}^N(t)-\mathbf{e}_i \in \chi\}} \\ \mathbf{Q}^N(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{e}_i) = \gamma_i Y_i^N(t) \frac{Y_0^N(t)}{N-1} + \\ \quad + \lambda_i Y_0^N(t) \mathbf{1}_{\{\mathbf{Y}^N(t)+\mathbf{e}_i \in \chi\}} \end{cases} . \quad (1)$$

Now, we show that the indicator functions in the transition rates of equation (1) are redundant, that is, if $\mathbf{Y}^N(0,\omega) \in \chi$ for all $\omega \in \Omega$, then $\mathbf{Y}^N(t,\omega) \in \chi$ for all $t > 0$ and $\omega \in \Omega$, regardless of the presence of these constraints. In this case, $\mathbf{Q}^N$ can be simplified to

$$\begin{cases} \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - \mathbf{e}_i\right) &= \mu_i Y_i^N(t) \\ \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{e}_i\right) &= \gamma_i Y_i^N(t) \frac{N - \mathbf{1}^\top \mathbf{Y}^N(t)}{N-1} + \\ & \quad + \lambda_i (N - \mathbf{1}^\top \mathbf{Y}^N(t)) \\ \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{v}\right) &= 0, \text{ if } \mathbf{v} \neq 0, \mathbf{v} \neq \pm \mathbf{e}_i \end{cases} . \quad (2)$$

Indeed, let $Y_i^N(t) = 0$. Equation (2) yields

$$\begin{cases} \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - \mathbf{e}_i\right) &= Q_i^N(0,-1) = 0 \\ \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{e}_i\right) &= \lambda_i \left(N - \sum_j Y_j^N(t)\right) \end{cases},$$

where $\mathbf{Q}_i^N$ is the rate transition matrix associated to $\left(Y_i^N(t)\right)$. Note that when $Y_i^N(t) = 0$, $\left(Y_i^N(t)\right)$ cannot decrease from 0 to -1 as conveyed in equation (2). In this case and if $Y_0^N(t) > 0$, the process $\left(Y_i^N(t)\right)$ can only increase due to external arrivals since there are no i-infected nodes in the network to spread the virus. Also, if $Y_i^N(t) = N - \sum_{j\neq i} Y_j^N(t)$ then, there is no more room for infections, and equation (2) yields

$$\begin{cases} \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) + \mathbf{e}_i\right) &= 0 \\ \mathbf{Q}^N\left(\mathbf{Y}^N(t), \mathbf{Y}^N(t) - \mathbf{e}_i\right) &= \mu_i Y_i^N(t) \end{cases} .$$

Otherwise, $\mathbf{1}_{\{\mathbf{Y}^N(t)-e_i\in\chi\}} = 1$ and $\mathbf{1}_{\{\mathbf{Y}^N(t)+e_i\in\chi\}} = 1$ and we conclude that in any of the cases, equation (1) boils down to (2).

C. Mean Field

We can now invoke Dynkin's formula (refer to section *III*.10 of Rogers and Williams [11]) and the characterization of $\mathbf{Q}^N$ developed in the previous section to deduce the stochastic integral equation underlying the infection dynam-

ics. Since $\left(Y_i^N(t)\right)$ is a Markov process,

$$\begin{aligned} M_i^N(t) &= Y_i^N(t) - Y_i^N(0) \\ &\quad - \int_0^t \sum_{\mathbf{p}\in\chi-\{\mathbf{Y}^N(s)\}} \mathbf{Q}\left(\mathbf{Y}^N(s),\mathbf{p}\right)\left(p_i - Y_i^N(s)\right) ds \\ &= Y_i^N(t) - Y_i^N(0) \\ &\quad - \int_0^t \mathbf{Q}\left(\mathbf{Y}^N(s), \mathbf{Y}^N(s) + \mathbf{e}_i\right) ds \\ &\quad + \int_0^t \mathbf{Q}\left(\mathbf{Y}^N(s), \mathbf{Y}^N(s) - \mathbf{e}_i\right) ds \end{aligned} \tag{3}$$

defines a martingale with respect to the natural filtration induced by the Poisson point processes associated to the arrivals, contamination, and healing processes. From equations (3) and (2), we have

$$\begin{aligned} Y_i^N(t) &= M_i^N(t) + Y_i^N(0) + \int_0^t \lambda_i \left(N - \mathbf{1}^\top \mathbf{Y}^N(s)\right) ds \\ &\quad + \int_0^t \gamma_i Y_i^N(s) \left(N - \mathbf{1}^\top \mathbf{Y}^N(s)\right)/(N-1) ds \\ &\quad - \int_0^t \mu_i Y_i^N(s) ds. \end{aligned} \tag{4}$$

Equation (4) defines a pathwise integral equation. That is, for a fixed $\omega \in \Omega$, $\left(Y_i^N(t,\omega)\right)_{t\geq 0}$ satisfies the integral equation (4). Remark that $\overline{Y}_i^N(t) = \frac{Y_i^N(t)}{N}$ is the fraction of i-infected nodes at time instant t with $\overline{\mathbf{Y}}^N(t) = \left[\overline{Y}_1^N(t)\, \overline{Y}_2^N(t) \ldots \overline{Y}_K^N(t)\right]^\top \in \Delta$, where $\Delta = \left\{\mathbf{y}\in\mathbb{R}^K : \mathbf{y}\geq 0, \mathbf{y}^\top \mathbf{1} \leq 1\right\}$. As N goes to infinity, the randomness underlying $\left(\overline{\mathbf{Y}}^N(t)\right)$ dies out, and it converges to the solution of a differential equation as stated in the following theorem.

Theorem 1. *If $\overline{\mathbf{Y}}^N(0) = \mathbf{z} \in \Delta$, then the sequence $\mathbf{Y}^N(t)$ converges in the Skorokhod topology to the solution $(\mathbf{y}(t))$ of the differential equation*

$$\frac{d}{dt} y_i(t) = \underbrace{\left(\lambda_i + \gamma_i y_i(t)\right)\left(1 - \mathbf{1}^\top \mathbf{y}(t)\right) - \mu_i y_i(t)}_{F_i(\mathbf{y}(t))} \tag{5}$$

for $i = 1, 2, \ldots, K$.

Proof: For reasons similar to reference [2], $\left(\overline{M}_i^N(t)\right)$ is a square integrable martingale and $\left(\overline{M}_i^N(t)\right)$ converges a.s. to 0 uniformly on compact sets. Similarly to Hunt and Kurtz [5], it is easy to check that $\left(\overline{Y}_i^N(t)\right)$ is asymptotically Lipschitz, and therefore it defines a relatively compact sequence (see also Ethier and Kurtz [9],) that is, there exists a subsequence $\left(\overline{Y}_i^{N_p}(t)\right)$ converging weakly to a limiting process $\left(\widetilde{Y}_i(t)\right)$. Provided that the martingale converges to zero and resorting to equation (4), the limiting process $\left(\widetilde{Y}_i(t)\right)$ satisfies

$$\begin{aligned} \widetilde{Y}_i(t) &= \widetilde{Y}_i(0) + \int_0^t \lambda_i \left(1 - \mathbf{1}^\top \widetilde{\mathbf{Y}}^N(s)\right) ds \\ &\quad + \int_0^t \gamma_i \widetilde{Y}_i^N(s) \left(1 - \mathbf{1}^\top \widetilde{\mathbf{Y}}^N(s)\right) ds \\ &\quad - \int_0^t \mu_i \widetilde{Y}_i^N(s) ds \end{aligned}$$

in a pathwise sense. Also, $\mathbf{F}$ (as defined in equation (5)) is globally Lipschitz over the simplex Δ, which implies that the initial value solution is unique (refer to V.I. Arnold [12],) and therefore $\left(\overline{\mathbf{Y}}^N(t)\right)$ converges weakly (endowing the trajectories space with the Skorokhod topology) to the solution $(\mathbf{y}(t))$, since the weak limit of any convergent subsequence $\left(\overline{Y}_i^{N_p}(t)\right)$ is $(\mathbf{y}(t))$.

V. Two Dimensional Qualitative Analysis

In this section, we analyze the qualitative behavior of the dynamical system $(\mathbf{y}(t))$ in theorem 1. The following discussion is similar (but not the same) to section 4.11 of [14] about the principle of competitive exclusion. We restrict ourselves to two classes of virus (for the sake of simplicity,) assuming $\lambda_i = 0$, $\mu_i > 0$ and $\gamma_i > 0$ for $i = 1, 2$, that is, there are no arrivals and the viral epidemic is driven purely through contamination from an initial infected population. The analysis as well as the conclusions presented in this section can be naturally extended to the general case of K classes of viruses. Our aim is to back up analytically the observed metastable behavior presented (via numerical simulations) in section VI. We start by identifying the singular (or equilibrium) points and specifying their stability nature – whether it is stable, saddle, center, etc. Clearly, the resulting singular points depend on the parameters μ_i and γ_i. An equilibrium point $\mathbf{y} \in \Delta$ is such that $F(\mathbf{y}) = 0$, where $F(\mathbf{y}) = [F_1(\mathbf{y})\, F_2(\mathbf{y})]^\top$ is defined in equation (5). Defining $y_0(t) = 1 - y_1(t) - y_2(t)$ as the fraction of healthy nodes, the Jacobian of the map F is given by

$$DF(y) = \begin{bmatrix} \gamma_1\left(y_0 - y_1 - \frac{\mu_1}{\gamma_1}\right) & -\gamma_1 y_1 \\ -\gamma_2 y_2 & \gamma_2\left(y_0 - y_2 - \frac{\mu_2}{\gamma_2}\right) \end{bmatrix}. \tag{6}$$

The equilibrium states are sensitive to the ratios $\tau_i = \frac{\gamma_i}{\mu_i}$. We now identify and classify the equilibrium points for different values of the effective transmission rates τ_1 and τ_2.

- $\tau_2 \leq \tau_1 < 1$: The equation $F(\mathbf{y}) = 0$ leads to

$$y_i y_0 - \frac{1}{\tau_i} y_i = 0 \Leftrightarrow y_i\left(y_0 - \frac{1}{\tau_i}\right) = 0.$$

If $\tau_i < 1$ (rate of healing greater than the rate of contamination for virus i) then, $y_i = 0$ since $0 \leq y_0 \leq 1$. Therefore, if $\tau_2 \leq \tau_1 < 1$, then the only equilibrium point is the origin $y = 0$. From the Jacobian (6) we conclude that $\mathbf{y} = 0$ is a stable equilibrium point since all the eigenvalues of $DF(0)$ are strictly negative real numbers. In words, the solution $(\mathbf{y}(t))$ converges to $\mathbf{y} = 0$ as t goes to infinity for an initial condition close enough to the origin. Moreover, the vector field $F(\mathbf{y}) < 0$ for all $y \in \overset{\circ}{\Delta}$, where we defined $x > 0$ if $x \in \mathbb{R}^n_{++}$. This indicates that solutions $(\mathbf{y}(t))$ are indeed attracted towards $\mathbf{y} = 0$ for any initial condition $\mathbf{y}(0) \in \overset{\circ}{\Delta}$, that is, if $\tau_1 < 1$ and $\tau_2 < 1$ then, the viruses will eventually die out.

- $\tau_2 < 1 < \tau_1$: It follows from the previous analysis that, if $\mathbf{y} \in \Delta$ is an equilibrium point, then $y_2 = 0$. Also

$$y_1\left(y_0 - \frac{1}{\tau_1}\right) = 0 \quad \Leftrightarrow \quad y_1 = 0 \text{ or } y_1 = 1 - \frac{1}{\tau_1}.$$

Therefore, we have two possible equilibrium states $\mathbf{y}^{(0)}=0$ and $\mathbf{y}^{(1)}=\left[1-\frac{1}{\tau_1}\ 0\right]^\top$ with $\mathbf{y}^{(0)}=0$ being a saddle point since $\left(1-\frac{1}{\tau_1}\right)>0$ and $\left(1-\frac{1}{\tau_2}\right)<0$. That is, the solution expands in the direction $\mathbf{e}_2$ about 0 and contracts in the direction $\mathbf{e}_1$ by future iterates. Moreover, since the Jacobian at $\mathbf{y}^{(1)}$ is a triangular matrix, its eigenvalues are given by the diagonal elements $\gamma_1\left(\frac{1}{\tau_1}-1\right)$ and $\gamma_2\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)$. Hence, the point $\mathbf{y}^{(1)}$ is a stable point and the solutions will converge to the equilibrium $\mathbf{y}^{(1)}$. Therefore, virus type 2 eventually dies and virus 1 survives.

- $1<\tau_2<\tau_1$: Through similar arguments, we can conclude that in this case $y_1=0$ or $y_1=1-\frac{1}{\tau_1}$, $y_2=0$ or $y_2=1-\frac{1}{\tau_2}$. The point $\mathbf{y}=[1-\frac{1}{\tau_1}\ 1-\frac{1}{\tau_2}]>0$ cannot be an equilibrium point. Indeed, in this case we would have $y_0=\frac{1}{\tau_1}=\frac{1}{\tau_2}$ which cannot hold true. The equilibrium points $\mathbf{y}^{(0)}=0$, $\mathbf{y}^{(1)}=\left[1-\frac{1}{\tau_1}\ 0\right]^\top$ and $\mathbf{y}^{(2)}=\left[0\ 1-\frac{1}{\tau_2}\right]^\top$ are unstable, stable, and saddle, respectively. We analyze if the basin of attraction associated to the stable point $\mathbf{y}^{(1)}$ comprises the whole set $\overset{\circ}{\Delta}$, that is, $\lim_{t\to\infty}\mathbf{y}(t)=\mathbf{y}^{(1)}$ independently of the initial condition $\mathbf{y}(0)\in\overset{\circ}{\Delta}$. In words, at most one type of virus survives (the one with greatest effective spreading rate) and the network is eventually stripped from the remaining species. We draw the vector field $\mathbf{F}$ for different regions of Δ. We can summarize in table I the phase portrait of the vector field $\mathbf{F}$. $\{S_i\}_{i=1,\dots,5}$ as defined in table I form

Region	$\mathbf{F}$
$S_1=\left\{\mathbf{y}\in\Delta: 1-y_1-y_2>\frac{1}{\tau_2}\right\}$	$\mathbf{F}>0$
$S_2=\left\{\mathbf{y}\in\Delta: \frac{1}{\tau_2}>1-y_1-y_2>\frac{1}{\tau_1}\right\}$	$\mathbf{F}_1>0, \mathbf{F}_2<0$
$S_3=\left\{\mathbf{y}\in\Delta: \frac{1}{\tau_1}>1-y_1-y_2\right\}$	$\mathbf{F}<0$
$S_4=\left\{\mathbf{y}\in\Delta: 1-y_1-y_2=\frac{1}{\tau_1}\right\}$	$\mathbf{F}_1$=0, $\mathbf{F}_2<0$
$S_5=\left\{\mathbf{y}\in\Delta: 1-y_1-y_2=\frac{1}{\tau_2}\right\}$	$\mathbf{F}_1>0, \mathbf{F}_2$=0

TABLE I
SIGN OF VECTOR FIELD $\mathbf{F}$ OVER $\{S_i\}$.

a partition for Δ. Figure 1 depicts in the plane the vector field $\mathbf{F}$. It also illustrates the partition $\{S_i\}$ of Δ. Set S_1 lies in the region between the dashed line and the axis. Set S_2 comprises the region between the dashed and dotted straight lines. Set S_3 is above the dotted straight line. Sets S_4 and S_5 represent the dotted and dashed straight lines, respectively. It suggests that any solution starting at $S_1\cup S_3\cup S_4\cup S_5$ is attracted towards S_2 (region between the dashed and dotted lines) and once it lies within the closure of S_2 it will remain there for all $t\geq 0$. Moreover, our dynamical system drifts within this region towards $\mathbf{y}^{(1)}=\left[1-\frac{1}{\tau_1}\ 0\right]$. The following simulation backs up our qualitative analysis. In figure 2 we present a numerical solution for different initial conditions. We included the straight lines S_4 and S_5 in figure 2 to stress that, in fact, all integral curves accumulate in S_2 and are further attracted towards $\mathbf{y}^{(1)}=[0\ 0.75]$. Also, Figure 2

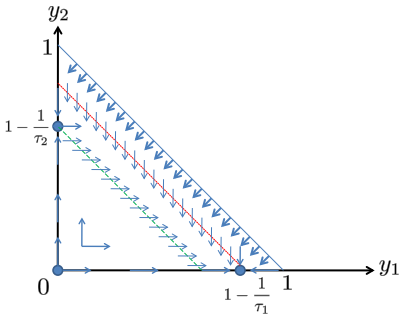

Fig. 1. Sketch of the vector field $\mathbf{F}$.

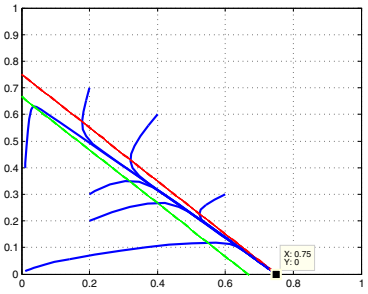

Fig. 2. Phase portrait of $\mathbf{F}$ (solid line). Parameters were set to $\gamma_1=4$, $\gamma_2=3$ and $\mu_1=\mu_2=1$.

suggests that the basin of attraction with respect to $\mathbf{y}^{(1)}$ is given by the whole set Δ. In light of this and assuming $\tau_i\gg 1$, we observe an underlying non-trivial metastable behavior: a slight dominance of virus type 1 on the effective spreading rate guarantees that the system drifts towards a highly unbalanced steady-state with $y_1=1-\frac{1}{\tau_1}\approx 1$ and $y_2=0$, independently of the initial conditions. Namely, it means that this system is quite sensitive to small perturbations on the space of parameters. In the next section we explore this behavior through numerical simulations. The strict case $\tau_i=1$ for some i bears no interest since it is lost under small perturbations.

VI. SIMULATION RESULTS

In this section we present numerical results that explicitly reveal the metastable behavior underlying the dynamical system $(\mathbf{y}(t))$. The system is much less sensitive regarding perturbations on the exogenous arrival rates λ_i as F depends linearly on this parameter. We assume that the epidemic spreads over the network due to intra-network contamination and exogenous infection. As we have seen, the equilibrium points depend on the statistical parameters through ratios and not exactly on the particular values of each parameter

$$\mathbf{F}(\mathbf{y})=0 \quad\Leftrightarrow\quad \frac{\lambda_i}{\mu_i}\left(1-\mathbf{1}^\top\mathbf{y}\right)+\frac{\gamma_i}{\mu_i}y_1\left(1-\mathbf{1}^\top\mathbf{y}\right)-y_1=0,\forall_i.$$

Therefore, we fix the healing and arrival rates to $\lambda_i=0.1$, $\mu_i=1$ for all $i=1,2,\dots,K$, and vary the transmission γ_i rate. In what follows, we consider 3 classes of viruses and draw the solution $(\mathbf{y}(t))$ of our mean field differential equation for

different values of γ_i. Figure 3 shows a sharp difference in the steady-state for the different classes. A small arrival λ_3 rate is enough to guarantee that virus 3 enters the network (since $y_3(0) = 0$) and infects a substantial fraction of the network. Figures 3 and 4 depict the evolution over time of

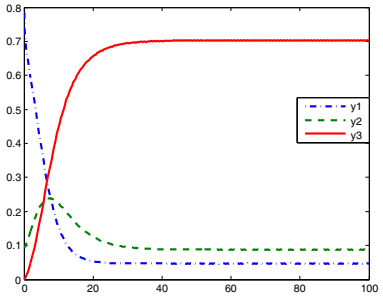

Fig. 3. Evolution of $(y_1(t))$, $(y_2(t))$ and $(y_3(t))$ for $t \geq 0$ with $y_1(0) = 0.8$, $y_2(0) = 0.1$ and $y_3(0) = 0$. We set $\gamma_1 = 4$, $\gamma_2 = 5$, $\gamma_3 = 6$ and $\lambda_i = 0.1$ for all $i = 1,2,3$.

the fractions $(y_1(t))$ (dash-dotted line,) $(y_2(t))$ (dashed line), and $(y_3(t))$ (solid line). For figure 3, we set $\gamma_1 = 4$, $\gamma_2 = 5$ and $\gamma_3 = 6$ whereas for figure 4 we increase the parameters to $\gamma_1 = 240$, $\gamma_2 = 250$ and $\gamma_3 = 260$. All figures present critical behavior and they illustrate that as τ_i increases and $\tau_i \gg 1$ this metastable behavior becomes even more significant. To sum

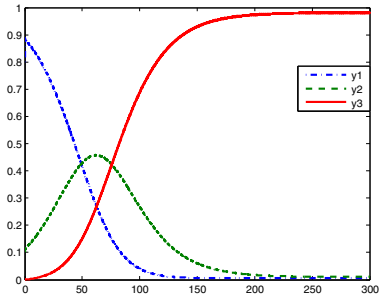

Fig. 4. Evolution of $(y_1(t))$, $(y_2(t))$ and $(y_3(t))$ for $t \geq 0$ with $y_1(0) = 0.8$, $y_2(0) = 0.1$ and $y_3(0) = 0$. We set $\gamma_1 = 240$, $\gamma_2 = 250$, $\gamma_3 = 260$ and $\lambda_i = 0.1$ for all $i = 1,2,3$.

up, we observe that, for larger ratios $\tau_i = \frac{\gamma_i}{\mu_i}$ (heavy-traffic), the system is more sensitive to small perturbations on the effective contamination rate τ_i. A small dominance of one of these parameters τ_i may be crucial for the contamination of a large fraction of the network with the dominant virus. Moreover, the system admits only one stable equilibrium point independent on initial conditions.

VII. Conclusion

We presented a microscopic model to study the dynamics of a multiple virus infection on a complete large network. The model tracks the evolution of the global state process $\left(\overline{Y}_i^N(t)\right)$: the fraction of i-infected nodes in the complete large network. We explored the mean field property of the system, namely, we proved that the stochastic process $\left(\overline{\mathbf{Y}}^N(t)\right)$ converges to the solution of a system of deterministic differential equations as the number of nodes N goes to infinity. In words, for a large network, the randomness on $\left(\overline{\mathbf{Y}}^N(t)\right)$ becomes arbitrarily negligible as N grows large. In the heavy-traffic environment, the asymptotic behavior of the system $(\mathbf{y}(t))$ is quite sensitive to perturbations in the parameter space, more precisely on the effective spreading rate $\tau_i = \frac{\gamma_i}{\mu_i}$ that underlies the speed of propagation of virus i. This means that a virus can suddenly occupy a large fraction of the network once it has a small advantage on the ratio τ_i over the other classes of viruses, even if it lags behind in population or it is close to extinction. The next natural step is to study virus propagation over networks with nontrivial topology, for instance, regular networks.

References

[1] Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibbi. " Stochastic Networks with Multiple Stable Points," *Annals of Applied Probability* Vol.36, No.1, 255-278, 2008.

[2] Nelson Antunes, Christine Fricker, Philippe Robert, and Danielle Tibbi. " Analisys of Loss Networks with Routing," *Annals of Applied Probability* 16 (2006), No.4, 2007-2026.

[3] Soummya Kar and José M. F. Moura. " Global Emergent Behaviors in Clouds of Agents," *ICASSP'2011 IEEE International Conference on Signal Processing*, Prague, May 2011. Invited Paper in Special Session on Bioinspired Signal Processing.

[4] Piet Van Mieghem, Jasmina Omic, and Robert Kooij. " Virus Spread in Networks," *IEEE/ACM Transaction on Networking*, Vol.17, No.1, pp.1-14, February (2009.)

[5] Hunt,P.J. and Kurtz,T.G. (1994.) " Large Loss Networks," *Stochastic Processes. Appl.* 53 363-378.

[6] Bailey, N.T.J. *The Mathematical Theory of Infectious Diseases*. London: Griffin, 1975.

[7] Mathew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

[8] Kermack W.O., and A.G. McKendrick. *A Contribution to the Mathematical Theory of Epidemics*. Proceedings of the Royal Society of London Series A 115 : 700–721.

[9] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.

[10] Phillipe Robert. *Stochastic Networks and Queues*. Springer-Verlag, 2003.

[11] L.C.G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*. Cambridge University Press, 2nd Edition, Volume 1, 2000.

[12] Vladimir I. Arnold. *Ordinary Differential Equations*. Springer-Verlag, 2nd Edition, 2006.

[13] Meyers, L.A. (2007). Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. Bulletin of the American Mathematical Society 44: 63-86.

[14] Martin Braun. *Differential Equations and Their Applications: An Introduction to Applied Mathematics*. Springer-Verlag, 3rd Edition, 1983.