# Noisy mean field game model for malware propagation in opportunistic networks

Hamidou Tembine, Vilanova Pedro, Merouane Debbah

# Noisy mean field game model for malware propagation in opportunistic networks

Hamidou Tembine, Pedro Vilanova, and Mérouane Debbah

Ecole Supérieure d'Electricité,
Chaire Alcatel-Lucent en radio-flexible,
Supelec, France
AMCS, KAUST, Saudi Arabia

**Abstract.** In this paper we present analytical mean field techniques that can be used to better understand the behavior of malware propagation in opportunistic large networks. We develop a modeling methodology based on stochastic mean field optimal control that is able to capture many aspects of the problem, especially the impact of the control and heterogeneity of the system on the spreading characteristics of malware. The stochastic large process characterizing the evolution of the total number of infected nodes is examined with a noisy mean field limit and compared to a deterministic one. The stochastic nature of the wireless environment make stochastic approaches more realistic for such types of networks. By introducing control strategies, we show that the fraction of infected nodes can be maintained below some threshold. In contrast to most of the existing results on mean field propagation models which focus on deterministic equations, we show that the mean field limit is stochastic if the second moment of the number of object transitions per time slot is unbounded with the size of the system. This allows us to compare one path of the fraction of infected nodes with the stochastic trajectory of its mean field limit. In order to take into account the heterogeneity of opportunistic networks, the analysis is extended to multiple types of nodes. Our numerical results show that the heterogeneity can help to stabilize the system. We verify the results through simulation showing how to obtain useful approximations in the case of very large systems.

## 1 Introduction

In modern times, the massive use of information formed an interconnected global society of billions in which communication systems are vital infrastructures. Among them, wireless communications is without doubt one of the most explosive developments ever to have taken place in the telecommunications industry. At the same time, over the last years, we have seen the power of microprocessors double about every 18 months, becoming considerably smaller, cheaper and abundant; indeed, they are ubiquitous and are even finding their way into everyday objects. Those technology trends - tiny, cheap processors with integrated sensors and wireless communications formed the technological basis for a new era of vast number of smart objects that communicate via wireless links.

Those incredibly diverse and complex wireless communication networks brought several fundamental technical issues: networks must be designed to carry out the intended functions in an efficient and predictable way; they must be manageable and upgradeable; and the most of all: reliable. Today, one of the major emerging threat against reliability is malware, that is, malicious self-replicating code. Threats posed by malware ranges from attacks against the confidentiality of the communication to attacks that actually alter the information traffic, hence destroying the integrity of the network.

Recently, malware outbreaks designed for personal computer environments on wired networks like those of Slammer and Code Red worms over the Internet have already inflicted severe economic damages, infecting thousands of hosts in short periods of time. Then, it is of vital importance to predict the limits of the damages that the attackers can inflict in large wireless networks, subject to its fundamental limitations, such as limited energy, unreliable communications and topology changes due mobility.

In this work we illustrate how mean field approaches can be used to reduce the complexity in the analysis of the damage that can be inflicted in a large network with opportunistic interaction of the objects. Opportunistic networking is one of the emerging communication paradigms in wireless mobile communications, in which the communication opportunities are based on sporadic and intermittent contacts. In contrast with the extensive work of malware propagation in wired networks ([1]), large wireless networks have not yet received similar attention, with a few exceptions ([2, 3]).

**Our contribution**

Contributions in this paper are twofold. First, from the theoretical point of view, we present a novel mean field approach, which tries to overcome one of the limitations of the "classical" mean field, that is, the approximation of an inherently stochastic system with a *deterministic* representation (ordinary differential equation) [4]. We propose a new, more general approach in which we preserve the main advantage of the classic mean field, that is, the reduction of the number of parameters in the analysis of large systems, but adding a random or "noisy" component. This new addition could lead to a more realistic mathematical model of the original situation, for example when many local object transitions occurs at the same time making the second moment unbounded. In this context the work [5] is not applicable anymore because the second moment of the number of object transitions per time slot may not vanish when the number of objects goes to infinity. A typical scenario is when many players do parallel transitions. The idea of the proposed analysis is that, if the third order in the Taylor approximation of the regular function of the mean field is bounded, then the noise may not be negligible but a convergence to a stochastic mean field limit can be established. Inspired from the work of [6] based on multidimensional diffusion processes, we were able to establish a mean field convergence to *non-deterministic* differential equations and extend the previous works in mean field interaction models (with and without controls). This new mean field limit which is stochastic is called

*"noisy" mean field limit* and applied in this work to malware propagation in opportunistic networks.

Second, from the malware propagation modelling point of view, we extend the model developed in [7] in which the types are not used and the impact of the control parameters are not specially studied. This leads to a limitation in the results obtained, because different types of systems could lead for example to slower rates of propagation. They can represent, for example, different operating systems, different versions of the operating systems or patched/unpatched version of the same operating system. To the best of our knowledge, in most of the related work about malware spreading in large networks authors do not model the heterogeneity of the network.

Most of the mean field studies do not examine consider control framework. We observe that control parameters are important in the mean field limit since they give new insights to uncontrolled mean field framework which may be constrained (energy limitation). This helps in controlling the proportion of infected nodes.

**Organization**

The rest of the paper is structured as follows. In next section we present basics on stochastic games with random set of interacting players. In section **??**, we overview existing mean field models in discrete time. We develop a controlled mean field framework in section 3 and in section 4 we provide a general convergence to mean field which is characterized by a stochastic differential equation and the payoff evolution are solution of partial differential equations. Finally, we apply the *noisy mean field* framework to opportunistic wireless large networks.

## 2 Basics of stochastic games with individual states

In this section we introduce basic notions of stochastic games and in subsequent sections we establish their convergence to the noisy mean field when the number of players goes to infinity under suitable conditions. A stochastic game with individual states and random set of interacting players is a collection

$$\Gamma = \left( \mathcal{N}, \mathcal{S}, (\mathcal{X}_j, \tilde{\mathcal{A}}_j, \mathcal{A}_j, r_j)_{j \in \mathcal{N}}, q, \mathcal{B}^n \right)$$

where
- $\mathcal{N}$ is the set of players. The cardinality of $\mathcal{N}$ is $n$.
- $\mathcal{S}$ is a set of environment states. A state vector.
- For every player $j$, $\mathcal{X}_j$ is its state space. The state of player $j$ at time $t$ is denoted by $X_j^n(t) = (\theta_j, Y_j^n(t))$ where $\theta_j$ is the type, which is constant through the game. The set of possible states $\mathcal{X}_j = \{1, 2, ..., \Theta\} \times \mathcal{Y}_j$ is finite. $\mathcal{Y}_j$ may include other parameters, such as, space location, current direction and so on.
- For every player $j \in \mathcal{N}$, $\tilde{\mathcal{A}}_j$ is the set of actions of that player. $\mathcal{A}_j : \mathcal{S} \times \mathcal{X}_j \longrightarrow 2^{\tilde{\mathcal{A}}_j}$ is a set-valued map (correspondence) that assigns to each state

$(s, x_j) \in \mathcal{S} \times \mathcal{X}_j$ the set of actions $\mathcal{A}_j(s, x)$ that are available to player $j$. We denote the state-action space by

$$\text{SXA} = \left\{ (s, x, a) : \; (s, x_1, \dots, x_n) \in \mathcal{S} \times \prod_{j \in \mathcal{N}} \mathcal{X}_j, \right.$$

$$\left. a = (a_j)_{j \in \mathcal{N}}, \; a_j \in \mathcal{A}_j(s, x), \; \forall \, j \in \mathcal{N} \right\}$$

• For every player $j \in \mathcal{N}$, $r_j : \; \text{SXA} \longrightarrow \mathbb{R}$ is an instant payoff function for player $j$.

• $q : \; \text{SXA} \longrightarrow \Delta(\mathcal{S} \times \prod_j \mathcal{X}_j)$ is a transition function where $\Delta(\mathcal{S} \times \prod_j \mathcal{X}_j)$ is the space of probability distributions over $\mathcal{S} \times \prod_j \mathcal{X}_j$. The marginal of $q$ is denoted by $q_j$, which represents the transition probabilities of state of player $j$, given the state of all other players and the resource state.

• $\mathcal{B}^n$ is a random set of players: $\mathcal{B}^n \subseteq \mathcal{N}$ denotes the set of interacting players at current time. The random selection will be explained in detail in the next section. The payoff function will be denoted by $r_j^{\mathcal{B}^n}$.

   *How does the dynamic game evolves?*

   The dynamic game starts at an initial $(s_0, x_0) \in \mathcal{S} \times \prod_j \mathcal{X}_j$ and is played as follows. Time space is $\mathbb{N}$. A set $\mathcal{B}_t^n$ of players are randomly selected without replacement for a one-shot interaction. The current state $s_t, x_t, \mathcal{B}_t^n$ is known by all the selected players. At each time $t \in \mathbb{N}$, each player $j \in \mathcal{B}_t^n$, chooses an action $a_{j,t} \in \mathcal{A}_j(s_t, x_t)$, and receives an instant payoff $r_j(s_t, x_t, a_t)$, where $a_t = (a_{j,t})_{j \in \mathcal{N}}$, and the game moves to a new state according to a probability distribution given by $q(. \mid s_t, x_t, a_t) \in \Delta(\mathcal{S} \times \prod_{j'} \mathcal{X}_{j'})$. The payoff and transitions of non-selected players are zero and the payoff function $r_j^{\mathcal{B}_t^n}$ of a selected player $j$ does not depend on the state-actions of the non-selected players.

**Strategies** Under complete information of past play, a history of length $t + 1$ corresponds to the sequence $h_t = (s_0, x_0, \mathcal{B}_0^n, a_0, s_1, x_1, \mathcal{B}_1^n, a_1, \dots, s_t, x_t, \mathcal{B}_t^n)$. A *pure* strategy $\sigma_j$ of player $j$ is a mapping that assigns to every finite history $h_t$ an element of $\mathcal{A}_j(s_t, x_t)$. A *mixed* strategy is a probability distribution of the pure strategies. The set of mixed strategies is denoted by $\Sigma_j$. A strategy profile $(\sigma_j)_{j \in \mathcal{N}}$ is *stationary* if $\forall \, j$, $\sigma_j(h_t)$ depends only on the current pair $(s_t, x_t)$. A stationary strategy of player $j$ can be identified with element of the product-space of $\prod_{(s,x) \in \mathcal{S} \times \mathcal{X}} \Delta(\mathcal{A}_j(s, x))$. Every profile $\sigma = (\sigma_j)_{j \in \mathcal{N}} \in \prod_j \Sigma_j$ of mixed strategies, together with the initial state $s_0, x_0$ induces a probability distribution $\mathbb{P}_{s_0, x_0, \sigma}$ over the space of infinite plays $\text{SXA}^\infty$. We denote the corresponding expectation operator by $\mathbb{E}_{s_0, x_0, \sigma}$.

**Payoff** We examine the finite-horizon payoff. Let $\sigma$ be a mixed strategy profile. In the finite-horizon payoff, a player considers the cumulative payoff during the first $T$ times. For every finite $T \in \mathbb{N}$, the finite-horizon payoff for player $j$ is

$$F_{j,T}(s_0, x_0, \sigma) = \mathbb{E}_{s_0, x_0, \sigma} \left[ \sum_{t=0}^{T-1} r_j^{\mathcal{B}_t^n}(s_t, x_t, a_t) \mathbb{1}_{\{j \in \mathcal{B}_t^n\}} \right] \tag{1}$$

where $\mathbb{1}_{\{\cdot\}}$ equals 1 if $\cdot$ is true and 0 if not.

## 3 Controlled mean field interaction model

In this section, we introduce a controlled mean field interaction model. The finite version of this model is a particular case of stochastic games with individual states. We restrict our attention into a particular class of behavioral strategies within we are able to establish the mean field convergence. This restriction is due to the fact that when the number goes to infinity, the dimension of the set of stationary strategies goes to infinity as well as. By letting the size of the system go to infinity, the discrete stochastic game problem is replaced by a limit of a system of Hamilton-Jacobi-Bellman equations coupled with a mean field limit ODE or coupled system of Bellman-Shapley optimality and discrete mean field evolution, that are deterministic and where the dimensionality of the original system has been transformed in the mass-behavior of the system.

Time $t \in \mathbb{N}$ is discrete. The *global state* of the system at time $t$ is $(S(t), X^n(t)) = (S(t), X_1^n(t), ..., X_n^n(t))$. Denote by $A^n(t) = (A_1^n(t), \ldots, A_n^n(t))$ the action profile at time $t$. The system $(S(t), X^n(t))$ is Markovian once the action profile $A^n(t)$ are drawn under Markovian strategies. We denote the set of Markovian strategies by $\mathcal{U}$. $M_x^n(t)$ is the fraction of players who belong to the population of individual state $x$. Similarly, we associate the process $U_a^n(t) = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{A_j^n(t)=a\}}$ to the fraction of actions.

**Strategies and random set of interacting players**

At time slot $t$, an ordered list $\mathcal{B}_t^n$, of players in $\{1, 2, \ldots, n\}$, is randomly selected without repetition as follows: First we draw a random number of players $k_t$ such that $\mathbb{P}(|\mathcal{B}_t^n| = k \mid M^n(t) = m) =: J_k^n(m)$ where the distribution $J_k^n(m)$ is given for any $n$, $m \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}^{|\mathcal{X}|}$. Second, we set $\mathcal{B}_t^n$ to an ordered list of $k_t$ players drawn uniformly at random among the $n(n-1)...(n-k_t+1)$ possible ones.

Each player such that $j \in \mathcal{B}_t^n$ takes part in a one-shot interaction at time $t$, as follows. First, each selected player $j \in \mathcal{B}_t^n$ chooses an action $a_{j,t} \in \mathcal{A}(s, x_j)$ with probability $u(a_j \mid s, x_j)$ where $(s, x_j)$ is the current player state. The stochastic array $u$ is the strategy profile of the population.

Denote the current set of interacting players $\mathcal{B}_t^n = \{j_1, \ldots, j_k\}$. Given the actions $a_{j_1}, ..., a_{j_k}$ drawn by the $k$ players, we draw a new set of individual states $(x'_{j_1}, ..., x'_{j_k})$ and resource state $s'$ with probability $L_{s;s'}^n(k, m, a)$, where $a$ is the vector of the selected actions by the interacting players.

We assume that for any given Markovian strategy, the transition kernel $L^n$ is invariant by any permutation of the index of the players within the same type.

This implies in particular that the players are only distinguishable through their individual state. Moreover, this means that the process $M^n(t)$ is also Markovian once the sequence of strategy is given. Denote by $w^n_{s,s'}(u, m)$ the marginal transition probability between the resource states. Given any Markov strategy and any vector $m$ of $\Delta$, the resource state generates an irreducible Markov decision process with limiting invariant measure $w_s(u, m)$. Then, we can simplify the analysis by fixing the resource state $S(t) = s$ without losing generality. The model is entirely determined by the probability distributions $J^n$, the transition kernels $L^n$ and the strategy profile $u$.

## 4 Noisy mean field approach

We provide a general convergence result of the mean field to a stochastic differential equation and a martingale problem is formulated for the the law of the process $M^n_t$. We are able to establish a mean field convergence to non-deterministic differential equations, thus, extending the previous works in mean field interaction, in mean field Markov decision teams, in mean field Markov games [5, 8, 9, 10]. We show that even if the expected number of players that do a transition in one time slot is not bounded, one can have a mean field limit, in such a case a stochastic one. This mean field limit is referred as *noisy mean field*.

Before presenting the main theoretical results of this paper, we first introduce some preliminary notions. The evolution of the system depends on the decision of the interacting players. Given a history $h_t = (S(0), X^n(0), A^n(0), \ldots, S(t) = s, X^n(t), A^n(t))$. $X^n(t+1)$ evolves according to the transition probability

$$L^n(x'; x, u, s) = \mathbb{P}\left(X^n(t+1) = x' \mid h_t\right)$$

The term $L^n(x'; x, u, s)$ is the transition kernel on $\mathcal{X}^n$ under the strategy $U^n$. Let $x^n = (x^n_1, \ldots, x^n_n)$ such that $\frac{1}{n} \sum_{j=1}^n \delta_{x^n_j} = m$ and define

$$\mathcal{L}^n(m'; m, u, s) := \sum_{\substack{(x'_1, \ldots, x'_n) \\ \frac{1}{n}\sum_{j=1}^n \delta_{x'_j} = m'}} L^n(x'; x, u, s)$$

$$= \mathbb{P}(M^n(t+1) = m' \mid M^n(t) = m, U^n(t) = u, S(t) = s) = \mathbb{P}(M^n(t+1) = m' \mid \tilde{h}_t)$$

where $\tilde{h}_t = (S(t'), X^n(t'), A^n(t'), t' \leq t, S(t) = s, X^n(t) = x^n)$. The term $\mathcal{L}^n(m'; m, u, s)$ corresponds to the projected kernel of $L^n$. Below we give sufficient conditions on the transition kernels to get a weak convergence of the process $M^n_t$ under the strategy $U^n(t)$.

We now present the main assumptions of this paper.
**A1:** $w_s(u, m)$ is continuously differentiable in $m$ and $u$. Note that this assumption was already implicit in the approach of [7] via the smoothness of the drift.
**A2:** There exists a continuous mapping $f : \mathbb{R}^{|\mathcal{M}|} \times \mathcal{U} \times \mathcal{S} \longrightarrow \mathbb{R}^{|\mathcal{M}|}$ such that $\forall s \in \mathcal{S}$;

$$\lim_n \sup_{u \in \mathcal{U}} \sup_{\|m\| \leq 1} \| \frac{f^n(m,u,s)}{\delta_n} - f(m,u,s) \| = 0$$

where

$$f_x^n(m,u,s) = \int_{m' \in \mathbb{R}^{|\mathcal{X}|}} \mathbb{1}_{\{\|m'-m\| \leq 2\}}(m_x' - m_x)\mathcal{L}^n(dm';m,u,s),$$

for $x \in \mathcal{X}$, and $s \in \mathcal{S}$. This assumption is analogous to H2 in which control parameters are added.

**A3:** There exists $\delta_n \searrow 0$ and a continuous mapping $a : \mathbb{R}^{|\mathcal{X}|} \times \mathcal{U} \times \mathcal{S} \longrightarrow \mathbb{R}^{|\mathcal{X}| \times |\mathcal{X}|}$ such that $\forall s \in \mathcal{S}$;

$$\lim_n \sup_{u \in \mathcal{U}} \sup_{\|m\| \leq 1} \| \frac{a^n(m,u,s)}{\delta_n} - a(m,u,s) \| = 0$$

where

$$a_{x,x',s}^n(m,u,s) =$$
$$\int_{m' \in \mathbb{R}^{|\mathcal{X}|}} \mathbb{1}_{\{\|m'-m\| \leq 2\}}(m_x' - m_x)(m_{x'}' - m_{x'})\mathcal{L}^n(dm';m,u,s),$$

for $(x,x',s) \in \mathcal{X}^2 \times \mathcal{S}$, where the third moment is finite. Note that under H3, $a \equiv 0$.

**A4:** For all $\epsilon > 0$ and $\forall s \in \mathcal{S}$;

$$\lim_n \sup_{u \in \mathcal{U}} \frac{1}{\delta_n} \int_{m' \in \mathbb{R}^{|\mathcal{X}|}} \mathbb{1}_{\{\|m'-m\| > \epsilon\}}\mathcal{L}^n(dm';m,u,s) = 0$$

Note that assumption H3 and lemma 3 from [7], implies $A4$.

**A4':** $\forall s \in \mathcal{S}$;

$$\sup_{u \in \mathcal{U}} \sup_{m \in \mathbb{R}^{|\mathcal{X}|}} \sup_{n \geq 1} \left[ \| \frac{a^n(m,u,s)}{\delta_n} \| + \| \frac{f^n(m,u,s)}{\delta_n} \| \right] < \infty$$

The smoothness assumptions H4-H5 in [7] imply $A4'$ which implies $A4$ too.

**Theorem 1** *Let $M_0^n \longrightarrow \pi$ in law where $\pi$ is a probability measure. Under A1-A4 [resp. A1-A4'], the process $M_t^n$ converges in law to a weak [resp. strong] solution of the Itô stochastic differential equation given by*

$$dm = \tilde{f}(m,u)dt + \tilde{\sigma}(m,u)dW_t$$

*where $\tilde{\sigma}\tilde{\sigma}^t = \tilde{a}$, $\tilde{f}(m,u) = \sum_s w_s(m,u)f(m,u,s)$, $\tilde{a}(m,u) = \sum_s w_s(m,u)a(m,u,s)$ and $w(m,u)$ is the invariant probability of the transition matrix $L(m,u)$*

A proof can be found in [11].

**Remark** This result generalizes the deterministic mean field limit conditions established in [7] for $\mathcal{U}$ equal to a singleton. It generalizes also the deterministic controlled mean field dynamics obtained in [5] for stationary strategies.

**Remark** The conditions $A1 - A4, A4'$ are weaker than those given in [7]. Under the conditions in [7], the noise term $\frac{1}{\delta_n}a^n \longrightarrow 0$ when $n$ goes the infinity. Moreover, the continuity assumption on the drift limit is not needed. If $\tilde{f}$ admits a unique integral curve, and $\tilde{a}$ is bounded and continuous then, the result applies as well. This allows us to apply it in wide range of networking scenarios with discontinuous drift limit but lower semi-continuity properties. Note that the uniformity in $u$ may not be satisfied. In that case, a local mean field solution is derived. An example of such discontinuity is provided in [5]. Our result extends also the convergence theorem in [5] which was restricted to stationary strategies. Here $u_t$ is an admissible strategy at time $t$.

**Definition 1 (Individual optimization framework)** *To a game as defined above, we associate a macroscopic population game, defined as follows. Each member $j$ of the population, with state $X_j(t)$ and a population profile $m[u](t)$. The initial condition of the game is $S(0) = s, X_j(0) = x,\ m[u](0) = m_0$. The population profile is solution to the noisy mean field evolution and $S(t), X_j(t)$ evolves as a jump process given by the marginal of $q$ which depends on $m(t),\ u(t)$ and the strategy $u'(t)$ of $j$. Further, let $V_T(s, x, u', u, m)$ be the $T-$stage payoff of player $j$ in this game, given that $S(0) = s, X_j(0) = x$ and $m(0) = m_0$, i.e. $V_T(s, x, u', u, m) =$*

$$\mathbb{E}\left( g(m(T)) + \int_0^T r(S(t), X_j(t), u'(t), u(t), m(t))dt \Big| s, x, m_0 \right)$$

*where $m$ is solution of the Itô SDE*

$$dm(t) = \tilde{f}(m(t), u(t))dt + \tilde{\sigma}(m(t), u(t))dW_t \qquad (2)$$

Using Itô's formula [12] the expected payoff evolution for a fixed horizon $T$ is then given by

$$\begin{cases} v_{u',u}(T, m) = g(m) \\ -\partial_t v_{u',u}(t, m) = \bar{r}(t, m, u, u') + \sum_{x \in \mathcal{X}} \tilde{f}_x(m, u)\partial_{m_x} v_{u',u}(t, m) \\ \qquad\qquad + \frac{1}{2} \sum_{(x,x') \in \mathcal{X}^2} \tilde{a}_{x,x'}(m, u)\partial^2_{m_x m_{x'}} v_{u',u}(t, m) \end{cases}$$

where $v_{u',u}(t, m) = \mathbb{E}_{S,X_j} V_T(s, x, u', u, m)$ starting at time $t$ and $\bar{r}(t, m, u, u') = \mathbb{E}_{S,X_j} r(S, X_j, u', u, m)$ is the global expected mean field payoff.

**Centralized mean field control:** We now provide the feedback optimality principle for the global expected mean field payoff $\bar{r}(t, m, u, u)$. Please note that in this case, each member of the population uses the same strategy. The $T-$stage mean field optimization problem writes

$$\sup_u \ \mathbb{E}g(m(T)) + \int_0^T \bar{r}(t, m(t), u(t), u(t))\ dt$$

subject to the stochastic differential equation (2).

A strategy $u_*(t) = \phi(t, m)$ constitutes an optimal mean field solution if there exist a continuously differentiable function $\bar{v}(t, m)$ satisfying the following Hamilton-Jacobi-Bellman equation combined with the mean field SDE. By combining the system, one gets the mean field optimality:

**Proposition 1** *The mean field optimality for horizon $T$ is given by*

$$
\begin{cases}
\bar{v}(T, m) = g(m) \\[2mm]
-\partial_t \bar{v} = \sup_{u_t \in \mathcal{U}_t} \left\{ \bar{r}(t, m_t, u_t, u_t) + \sum_{x \in \mathcal{X}} \tilde{f}_x(m_t, u_t) \partial_{m_x} \bar{v} \right. \\[2mm]
\qquad\qquad \left. + \frac{1}{2} \sum_{(x, x') \in \mathcal{X}^2} \tilde{a}_{x,x'}(m_t, u_t) \partial^2_{m_x m_{x'}} \bar{v} \right\} \\[2mm]
dm_t = \tilde{f}(m_t, u_t) dt + \tilde{\sigma}(m_t, u_t) dW_t, \ t > 0 \\[2mm]
m_0 = m
\end{cases}
$$

*where $\bar{v} = \bar{v}(t, m)$.*

Please note that this kind of systems may not have a solution. Counterexamples can be found in [11].

### Discrete time mean field limit

If the step-size $\delta_n$ is not vanishing when $n$ goes to infinity, then the deterministic mean field limit is in discrete time and driven by the probability transition $L_t(u, m)$. Given an initial population profile $m_0$ and a terminal payoff, the sequence of population profiles $\{m_t\}_t$ is driven by the transition probabilities $\{L_{t,s,x,s',x'}(u, m_t)\}_t$.

$$
m_{t+1}(x) = \sum_{x' \in \mathcal{X}} m_t(x') L_{t,x,x'}(u_t, m_t), \tag{3}
$$

where

$$
L_{t,x,x'}(u, m) = \sum_{k \geq 0} \sum_{s} w_s(u, m) L_{t,s,x,x'}(u, m; k) J_k(m),
$$

$L_{t,s,x,x'}(u, m; k)$ is the limiting probability transition from $x$ to $x'$ when the resource state is $s$ and the number of interacting players is $k$. Combining with the Bellman-Shapley optimality criterion, one gets the following system in the finite horizon case:

$$
\begin{cases}
v_{j,t}(s,x,m) = \max_{u_j} \left\{ r(s,x,u,m_t) \right. \\[1em]
\qquad\qquad \left. + \sum_{s',x'} L(s',x'|s,x,u,m_t)v_{j,t+1}(s',x') \right\} \\[1em]
m_{t+1}(x) = \sum_{x'\in\mathcal{X}} m_t(x')L_{t,x',x}(u,m_t)
\end{cases} \qquad (4)
$$

## 5 Application to malware propagation

In this section, we apply the mean field approach to a controlled malware propagation in opportunistic networks. The malware propagation model is based on [7] in which the impact of the control parameters is not examined and the player types are not used. The types can represent, for example, different operating systems, different versions of the operating systems, or patched/unpatched version of the same operating system. In most of the related work about malware spreading in large networks authors do not model the heterogeneity of systems which forms the network, as far as we know. This leads to a limitation in the results obtained, because different types of systems could lead for example to slower rates of propagation.

In this example we have mobile nodes that can be infected by a malicious code. There are two infected states: passive and active. Non infected nodes are susceptible. Then, the set of possible state of a node is $\{P,A,S\}$ (for passive, active and susceptible) and the set of possible types is $\{\theta_1,\theta_2\}$. The state of the system at time $t$ is $X^n(t) = (P_1(t), P_2(t), A_1(t), A_2(t), S(t))$, where $P(t) + A(t)+S(t) = n$, $\sum_j P_j(t) = P(t)$, $\sum_j A_j(t) = A(t)$, $\forall t$ and $n$ is the total number of mobiles in the system. In this example, there is no resource. The occupancy measure is $M_\theta^n(t) = (P_\theta(t)/n, A_\theta(t)/n, S(t)/n) = (P_\theta^n(t), A_\theta^n(t), S^n(t))$. At every time step we want to control the proportion of infected nodes, which is $I^n(t) := A^n(t) + P^n(t)$. There are two fundamental ways to get infected:

1. Caused by a system flaw. (e.g: an exploit that could allow arbitrary code execution).
2. Caused by human flaw. (e.g: the user is deceived and executes a dangerous piece of code).

We can model this system as a controlled mean field interaction model. The interaction is simulated using the following rules:

1. A passive node may become susceptible (inoculation) with probability $\delta_P$.
2. A passive node with type $\theta$ may opportunistically encounter another passive node of type $\theta'$, and both become active. This occurs with probability proportional to the frequency of other passive nodes at time $t$. For type $\theta$, the probability is $\lambda(P_{\theta'}^n(t) - \frac{1}{n}\mathbb{1}_{\{\theta=\theta'\}})$. Note that the passive node can decide to contact the other passive node or not, so there are two possible

actions: $\{m, \bar{m}\}$ (for meet and not meet). Those events will be modeled as a Bernoulli random variable with success (meeting) probability $\delta_m$, which represents $u(m|P, \theta)$. Here we model the possibility of getting infected by a system flaw.

3. An active node may become susceptible (inoculation) with probability $\delta_A$.

4. An active node of type $\theta$ may become passive with probability $\beta \frac{P_\theta^n(t)}{h_\theta + P_\theta^n(t)}$ at time $t$. Here is assumed that, at high concentrations of passive nodes, each active node infects some maximum number of passive ones per time step. This reflects finite total bandwidth. The parameter $0 \leq \beta \leq 1$ has the interpretation of the maximum infection rate. The parameter $0 \leq h_\theta \leq 1$ is the passive node density at which the infection proceeds at half of its maximum rate. Here we model the possibility of getting infected by a system flaw.

5. A susceptible node may become active with probability $\delta_S$

6. A susceptible node may become passive via two ways. First, $\delta_{Sm}$ is the probability of getting infected by a human flaw. In this case, the susceptible node can "decide" to get deceived or not, so there are two possible actions: $\{o, \bar{o}\}$. The stationary strategy in this case will be modeled as a coin toss with probability $\delta_e$. Second, $\eta(P_\theta^n(t) + P_{\theta'}^n(t))$ models the probability of encountering a passive node. In this case, the passive node can decide to contact the susceptible node or not, and it is modeled analogously to the other two cases.

At every time step, one of the transitions is randomly selected and performed. The number of nodes that do a transition in one time slot is always 0, 1 or 2. In order to control the infected population, each transition has a certain payoff contribution which could be 0 if no infected node is inoculated, $1/n$ if there is a node which is inoculated and $-1/n$ if one node is infected. In Table 1 are the transition probabilities, the contribution to $M^n(t+1) - M^n(t)$, the set of actions, and the contribution to the total payoff.

| Case | Transition proba. $(\theta, \theta' \in \{1,2\})$. | $M_\theta^n(t+1) - M_\theta^n(t)$ | Actions | Payoff contrib. |
|---|---|---|---|---|
| 1 | $P_\theta^n(t)\delta_P$ | $(-1, 0, 1)/n$ | singleton set | $-1/n$ |
| 2 | $P_\theta^n(t)\delta_m^2 \lambda(P_\theta^n(t) - \frac{1}{n})$ | $(-2, 2, 0)/n$ | $\{m, \bar{m}\}$ | $0$ |
| 3 | $A_\theta^n(t)\delta_A$ | $(0, -1, 1)/n$ | singleton set | $-1/n$ |
| 4 | $A_\theta^n(t)\beta \frac{P_\theta^n(t)}{h_\theta + P_\theta^n(t)}$ | $(-1, 1, 0)/n$ | singleton set | $0$ |
| 5 | $S^n(t)\delta_S$ | $(0, 1, -1)/n$ | singleton set | $1/n$ |
| 6 | $S^n(t)(\delta_e\delta_{Sm} + \delta_m\eta P^n(t))$ | $(1, 0, -1)/n$ | $\{o, \bar{o}, m, \bar{m}\}$ | $1/n$ |

**Table 1.** Probabilities, effects, actions and payoffs.

The intensity, that is, the probability that one arbitrary object does a transition in one time slot is of the order of $1/n$. The drift, that is, the expected change of $M^n$ in one time step, given the current state of the system is:

$$f_\theta^n(m) = n\mathbb{E}(M_\theta^n(t+1) - M_\theta^n(t)|M^n(t) = m) =$$

$$\begin{pmatrix} -p_\theta\delta_P - 2p_\theta\delta_m^2\lambda\frac{np_\theta-1}{n} - a_\theta\beta\frac{p_\theta}{h_\theta+p_\theta} + s(\delta_e\delta_{Sm}+\delta_m\eta(p_\theta+p_{\theta'})) \\ 2p_\theta\delta_m^2\lambda\frac{np_\theta-1}{n} - a_\theta\delta_A + a_\theta\beta\frac{p_\theta}{h_\theta+p_\theta} + s\delta_S \\ p_\theta\delta_P + a_\theta\delta_A - s\delta_S - s(\delta_e\delta_{Sm}+\delta_m\eta(p_\theta+p_{\theta'})) \end{pmatrix}$$

where $m = (p_\theta, p_{\theta'}, a_\theta, a_{\theta'}, s)$. Then the limit $f(m)$ is

$$\begin{pmatrix} -p_\theta\delta_P - 2\lambda p_\theta^2\delta_m^2 - a_\theta\beta\frac{p_\theta}{h_\theta+p_\theta} + s(\delta_e\delta_{Sm}+\delta_m\eta(p_\theta+p_{\theta'})) \\ -p_{\theta'}\delta_P - 2\lambda p_{\theta'}^2\delta_m^2 - a_{\theta'}\beta\frac{p_{\theta'}}{h_{\theta'}+p_{\theta'}} + s(\delta_e\delta_{Sm}+\delta_m\eta(p_\theta+p_{\theta'})) \\ 2\lambda p_\theta^2\delta_m^2 - a_\theta\delta_A + a_\theta\beta\frac{p_\theta}{h_\theta+p_\theta} + s\delta_S \\ 2\lambda p_{\theta'}^2\delta_m^2 - a_{\theta'}\delta_A + a_{\theta'}\beta\frac{p_{\theta'}}{h_{\theta'}+p_{\theta'}} + s\delta_S \\ (p_\theta+p_{\theta'})\delta_P + (a_\theta+a_{\theta'})\delta_A - 2s\delta_S - 2s(\delta_e\delta_{Sm}+\delta_m\eta(p_\theta+p_{\theta'})) \end{pmatrix}$$

### 5.1 Homogeneous system

We briefly mention the homogeneous mean field. The drift is obtained by computing the expected changes in one time slot: $f^n(m) = \mathbb{E}(M^n(t+1) - M^n(t)|M^n(t) = m) =$

$$\frac{1}{n}\begin{pmatrix} -p\delta_P - 2p\delta_m^2\lambda\frac{np-1}{n} - a\beta\frac{p}{h+p} + s(\delta_e\delta_{Sm} + \delta_m\eta p) \\ 2p\delta_m^2\lambda\frac{np-1}{n} - a\delta_A + a\beta\frac{p}{h+p} + s\delta_S \\ p\delta_P + a\delta_A - s\delta_S - s(\delta_e\delta_{Sm} + \delta_m\eta p) \end{pmatrix}$$

where $m = (p, a, s)$. Then the limit is

$$f(m) = \begin{pmatrix} -p\delta_P - 2p^2\delta_m^2\lambda - a\beta\frac{p}{h+p} + s(\delta_e\delta_{Sm} + \delta_m\eta p) \\ 2p^2\delta_m^2\lambda - a\delta_A + a\beta\frac{p}{h+p} + s\delta_S \\ p\delta_P + a\delta_A - s\delta_S - s(\delta_e\delta_{Sm} + \delta_m\eta p) \end{pmatrix}$$
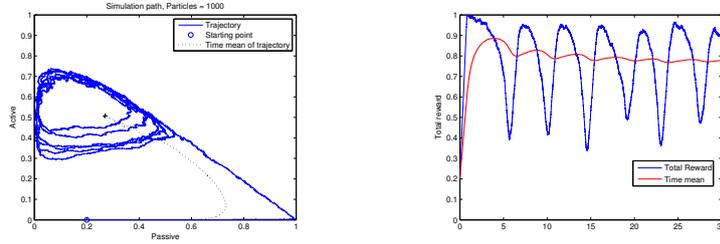
In all the simulations, we kept this parameters unchanged: $\beta = 10^{-2}$, $\delta_A = 5 \times 10^{-3}$ and $\delta_P = \delta_S = \delta_{Sm} = 10^{-4}$. On parameter $h$ (the passive node density at which the infection of active nodes proceeds at half of it maximum value) depends the stability of the system. Here we set $h = 10^{-2}$ in order to obtain an unstable behaviour. Regarding the control parameters, $\delta_m = 1$, $\delta_e = 1$ means no control.

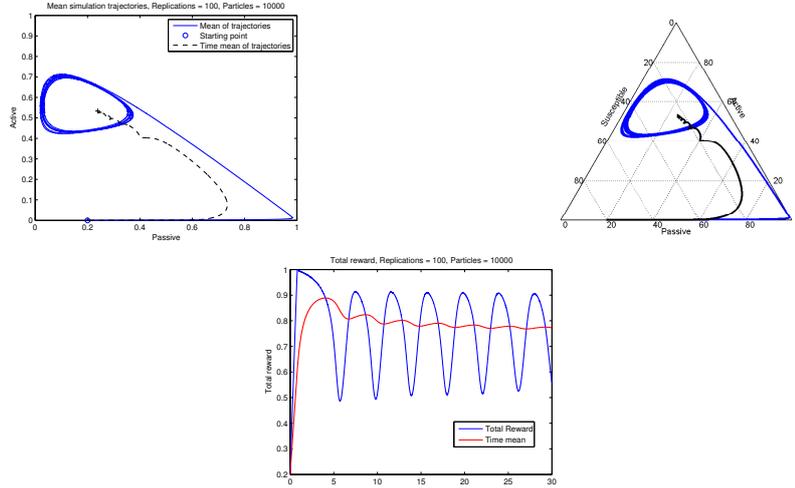We investigate the evolution of the system in the following scenarios:

– Trajectory of one run of the simulation (figure 1),
– Mean trajectory of multiple simulations (figure 2),
– Mean field limit trajectory (figure 3),
– Trajectory of the payoff function,
– Controlled mean field limit (figure 4),
– Trajectory of the heterogeneous malware propagation (figure 6),
– Optimal control under the mean field limit (figure 7),
– Noisy mean field (stochastic path), (figure 8).

These configurations are analyzed with control and without control parameters. We observe that the time mean $\frac{1}{T}\int_0^T m(s)ds$ is asymptotically close to the stationary point inside the limit cycle.

**Uncontrolled behaviour** In figures 1,2,3 we can see the simulation results, using the well-known algorithm for exact simulation of a discrete time Markov chain. The initial configuration is $(0.2, 0, 0.8)$. An oscillating behaviour of the total reward can be seen.



**Fig. 1.** Left: Path of one simulation without control. Right: Total reward.



**Fig. 2.** Leftmost: Mean of simulation trajectories without control. Rightmost: Total reward.

**Controlled behaviour** In order to illustrate the control parameters, suppose we want to keep the proportion of infected nodes below 0.9 for all times. One simple way to achieve this is to reduce the contact tendency of a passive node. Then, we set $\delta_m = 0.075$, which is the more relevant control parameter. The results can be seen in figures 4 and 5.

**Fig. 3.** Left: Deterministic mean field without control. Right: Total reward.
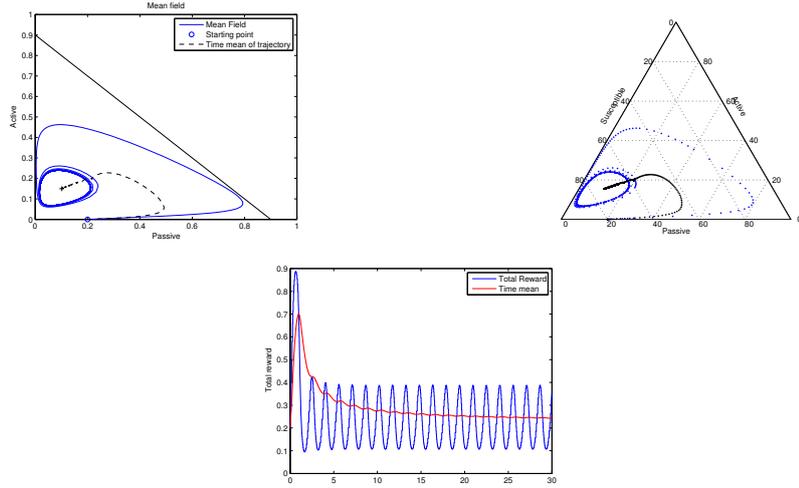


**Fig. 4.** Leftmost: Deterministic mean field with controlled passive node contact tendency. Rightmost: Total reward.

### 5.2 Heterogeneous system

We investigate numerically the behavior of the mean field limit for two types $\theta$ and $\theta'$. In figure 6 we can see that it is possible to stabilize the homogeneous system using classes.

### 5.3 Optimal strategy for the homogeneous system

Since the payoff function is the same for all the players i.e. $r_j(\cdot) = r(\cdot)$ the discounted stochastic game with common payoff can be transformed in a team problem. Moreover, the set of actions is the same for all the players. In figure 7, we plot the optimal strategy obtained by solving the system in the finite horizon case (equation (4)). The existence of a dominant strategy can be observed in the plot.
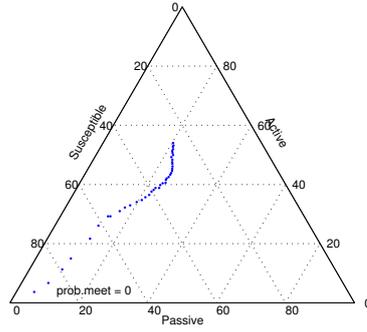
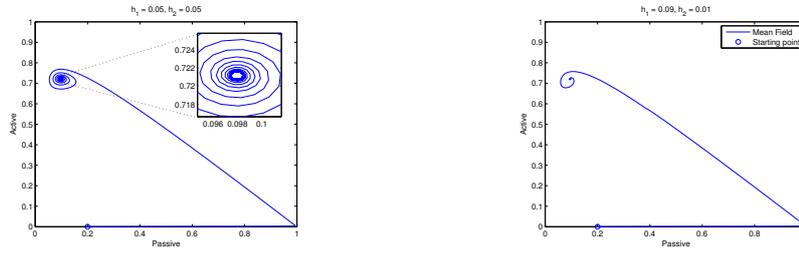**Fig. 5.** Evolution of the limit point of the time mean, for $\delta_m = 0, 0.025, .., 0.925, 1$.



**Fig. 6.** Left: Unstable deterministic mean field, for $h_1 = 0.05$, $h_2 = 0.05$. Right: Stable mean field $h_1 = 0.09$, $h_2 = 0.01$.
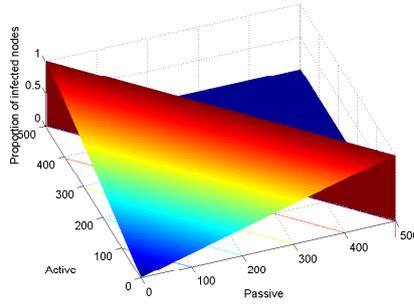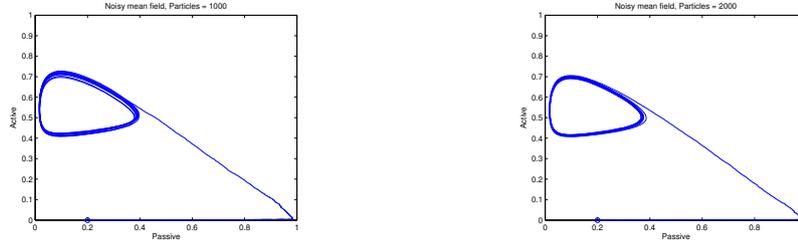


**Fig. 7.** Optimal strategy using backward induction for the homogeneous system.
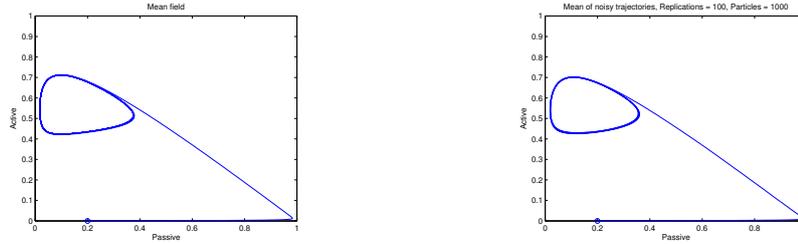
### 5.4 Noisy mean field

In order to give some feeling on how the noisy mean field evolves with time, we show in figure 8 two different realizations, for $n = 1000$ and $n = 2000$. The variance of the noise is obtained from $\tilde{\sigma}^n = \frac{1}{\delta_n} f^n - f$ which its norm sup is

bounded by $\frac{2D\lambda}{n}$, then we have $d\tilde{m} = f(\tilde{m})dt + \tilde{\sigma}^n dW_t$. Note that the smooth version of the last equation can be $\dot{m}^n = \frac{f^n(m^n)}{\delta_n}$. It is worth to mention that, in this case, the simulation algorithm is only exact when the numerical time step vanishes. In figure 9 we compare the deterministic mean field versus the mean



**Fig. 8.** Left: Path of one simulation of the noisy mean field, for $n = 1000$, Right: $n = 2000$.

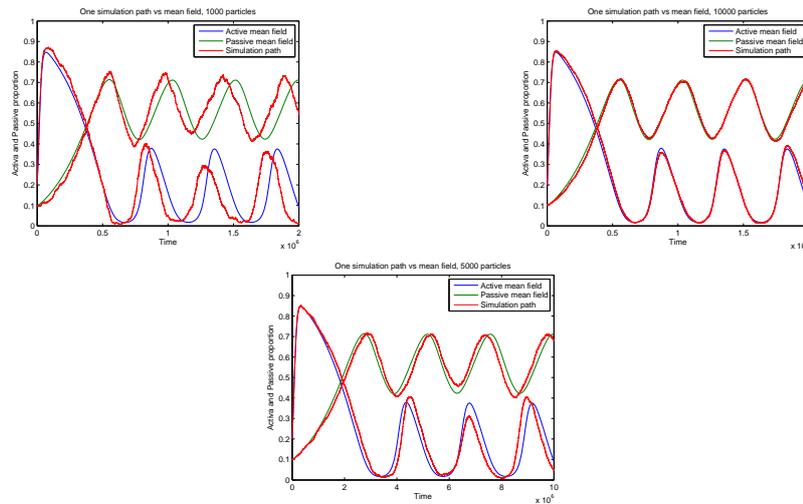trajectory of the noisy mean field.



**Fig. 9.** Left: Deterministic mean field, Right: Mean of noisy trajectories.

### 5.5 Effect of the system size

In this subsection we examine the effect of $n$ of the mean field object $M^n$. We represent the evolution (in time) for different value of $n$ and same state-component (see figures 10).

## 6 Concluding remarks

We have studied mean field stochastic games and established a mean field convergence to Itô stochastic differential equations. Our convergence result opens new questions for discontinuous mapping (namely the drift and the variance). In

**Fig. 10.** Mean field vs Simulation: Impact of $n$.

that case one may expect to have a stochastic differential inclusion as in [13] for which the existence of solution needs more attention. This leads to ask what will happen when the discontinuities are coming from the strategies as it is the case for the best response correspondence? We do not have answers to this question and postpone it to future work. In practical scenarios, when the second moment conditions are not satisfied, the conditions for mean field convergence given in [5, 14] are not applicable. In these cases, the weak conditions that we provided here cover a wide range of networking scenarios. The cycling behavior of the fraction of infected is sensitive to the control parameters. It would be useful to understand this relationship better by taking into account the network topology.

# References

1. C. C. Zou, W. Gong, D. F. Towsley, and L. Gao, "The monitoring and early detection of internet worms," *IEEE/ACM Trans. Netw.*, vol. 13, no. 5, pp. 961–974, 2005.
2. V. Karyotis, A. Kakalis, and S. Papavassiliou, "Malware-propagative mobile ad hoc networks: Asymptotic behavior analysis," *J. Comput. Sci. Technol.*, vol. 23, no. 3, pp. 389–399, 2008.
3. Y. Xu and W. Wang, "The speed of information propagation in large wireless networks," *IEEE Infocom*, 2008.
4. H. Tembine, E. Altman, R. ElAzouzi, and Y. Hayel, "Evolutionary games in wireless networks," *IEEE Trans. on Systems, Man, and Cybernetics, Part B, Special Issue on Game Theory*, vol. 40, pp. 634–646, June 2010.
5. H. Tembine, J. Y. Le Boudec, R. ElAzouzi, and E. Altman, "Mean field asymptotic of markov decision evolutionary games and teams," *in the Proc. of GameNets*, May 2009.

6. D. W. Stroock and S. R. S. Varadhan, "Multidimensional diffusion processes," *Springer Verlag, New York*, 1979.
7. M. Benaïm and J.-Y. Le Boudec, "A class of mean field interaction models for computer and communication systems," *Perform. Eval.*, vol. 65, no. 11-12, pp. 823–838, 2008.
8. G. Y. Weintraub, L. Benkard, and B. Van Roy, "Oblivious equilibrium: A mean field approximation for large-scale dynamic games," *Advances in Neural Information Processing Systems*, vol. 18, 2006.
9. H. Tembine, "Mean field stochastic games: convergence, q/h learning, optimality," *American Control Conference, ACC 2011, San Francisco, California, US*, 2011.
10. ——, "Hybrid mean field game dynamics in large populations," *American Control Conference, ACC 2011, San Francisco, California, US*, 2011.
11. ——, "Mean field stochastic games: Simulation, dynamics and applications," *Suplec*, 2010.
12. B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications (Universitext)*, 6th ed.   Springer, 7 2003.
13. M. Benaim, J. Hofbauer, and S. Sorin, "Stochastic approximation and differential inclusions," *SIAM J. Control and Optimization*, vol. 44, pp. 328–348, 2005.
14. H. Tembine, "Population games in large-scale networks: time delays, mean field and applications," *LAP Lambert Academic Publishing, ISBN 978-3-8383-6392-9, 250 pages*, December 2010.