SELF-ADAPTIVE CONGESTION CONTROL FOR MULTI-CLASS INTERMITTENT CONNECTIONS IN A COMMUNICATION NETWORK

CARL GRAHAM AND PHILIPPE ROBERT

Abstract. A Markovian model of the evolution of intermittent connections of various classes in a communication network is established and investigated. Any connection evolves in a way which depends only on its class and the state of the network, in particular as to the route it uses among a subset of the network nodes. It can be either active (ON) when it is transmitting data along its route, or idle (OFF). The congestion of a given node is defined as a functional of the transmission rates of all ON connections going through it, and causes losses and delays to these connections. In order to control this, the ON connections self-adaptively vary their transmission rate in TCP-like fashion. The connections interact through this feedback loop. A Markovian model is provided by the states (OFF, or ON with some transmission rate) of the connections. The number of connections in each class being potentially huge, a mean-field limit result is proved with an appropriate scaling so as to reduce the dimensionality. In the limit, the evolution of the states of the connections can be represented by a non-linear system of stochastic differential equations, of dimension the number of classes. Additionally, it is shown that the corresponding stationary distribution can be expressed by the solution of a fixed-point equation of finite dimension.

1. Introduction

The Internet network can be described as a very large distributed system, which manages the dynamic exchange of data transmission through connections established between nodes of the network under a very dynamic, random, scheme. Nodes cannot cope at all times with the huge amount of data transmitted through them by the varying connections, and congestion events occur, causing losses and delays to the connections.

AIMD algorithms (Additive Increase, Multiplicative Decrease) in the TCP protocol regulate Internet traffic: a connection increases gradually its throughput as long as congestion does not occur along its route (Additive Increase), but decreases it brutally when such an event takes place, usually by cutting it in half (Multiplicative Decrease).

An important practical issue is to obtain a better qualitative and quantitative understanding of such algorithms, and devising and analyzing pertinent mathematical models may be very helpful. There are many serious challenges, see Graham and Robert [7] for a discussion of the literature in this domain. Some objectives are to:

(1) Propose a stochastic model for the arrivals and durations of connections.
(2) Propose a stochastic model of flow control by AIMD algorithms; see, *e.g.*, Dumas *et al.* [5] and Guillemin *et al.* [9].
(3) Describe the state of the network, including the large number of connections interacting at its nodes through the congestion they create, in a mathematically tractable way.

Date: August 20, 2010.
Key words and phrases. Flow control algorithm and Multi-class system and Mean-field limit and Chaoticity and Nonlinear stochastic differential equation and Stationary distribution.

Let us be more specific. A client at a node initiates a connection to a server at another node, in order to transmit data for some specific purpose. Each data packet is routed from source to destination through intermediate nodes, using local routing decisions depending on the instantaneous state of the network. The connection adapts to the congestion encountered by its packets by adjusting its throughput as above, and ends once all the data has been completely transferred.

The same client may well repeat these steps at various instances with the same server, purpose, and reaction characteristics to network conditions, and thus intermittently establish connections of the same kind. For simplicity of expression, we call "user" such a client-server-purpose-reaction combination.

Graham and Robert [7] introduced and analyzed a Markovian network model, further studied in Graham *et al.* [8], constituted of J nodes and hosting K classes of permanent connections. This paper extends this to K classes of users establishing intermittent connections in an ON-OFF pattern. An OFF user is idle, and switches ON after a random duration. While ON, the user transmits under an AIMD scheme (after establishing a connection) and switches OFF (cutting the connection) after a random duration depending on throughput (related to node load). In particular, the number of on-going connections varies with time.

This gives a possible answer to Point 1 above; ON-OFF users are quite natural in models for data transmission in the Internet, for instance a web transaction (http connection) can be thought of as a succession of file transfers between server and client. For Point 2, the load of a node may be defined as a linear functional of the throughputs of all connections going through it, and the loss rate as a function of these quantities, as in [7].

Point 3 is achieved by proving a mean-field limit result, propagation of chaos, which we describe roughly. When the numbers of users in each class converge to infinity with limit ratios, for adequate initial conditions, asymptotically users evolve independently, and the limit behavior of users of class k for $1 \leq k \leq K$ is described by a nonlinear (McKean-Vlasov) stochastic differential equation (SDE) in $\mathbb{R}_+^K$, the effect of the interactions on the connections being encoded in the non-linearity of the SDE, as in [7].

The introduction of intermittent connections significantly changes the problems investigated in [7, 8]. A cemetery state -1 is added to encode the fact that a user can be OFF, and notably creates technical problems concerning the mandatory Lipschitz properties for mean-field convergence. Moreover, the limit stationary distributions are still characterized by finite vectors solving a fixed point equation, but their expression is much more complicated than before; it is related to the resolvent of a generic one-dimensional Markov process.

Section 2 introduces more precisely the stochastic model of the network, and the mean-field scaling. Section 3 gives the main properties of the limiting nonlinear SDE. Section 4 establishes the mean-field result. Section 5 studies equilibrium properties of the limit.

Acknowledgements
This work has been presented at the Stochastic Networks workshop in Cambridge in March 2010 in the programme on communication networks at the Newton Institute. We are grateful to the organizers for the invitation.

2. Markovian multi-class model for the network

2.1. **Classes of ON-OFF users, throughput, node load.** Recall that "user" denotes a possible kind of network use by a source-destination pair, and hence corresponds to the behavior, with respect to congestion, of its packet routes, rate of increase of throughput, drop rate, etc.

The stochastic network is constituted of $J \geq 1$ nodes and hosts $K \geq 1$ classes of users. There are $N_k \geq 1$ users of class k for $1 \leq k \leq K$, which bounds the *varying* number of possible on-going transmissions of this class. The notations

$$N = (N_1, \dots, N_K)\,, \qquad |N| = N_1 + \cdots + N_K\,,$$

will be used, and $|N|$ is the total number of users in the network.

Users alternate between being *active*, or ON, and *inactive*, or OFF. An ON user has data to transmit, and establishes a connection in the network, with throughput controlled by a dynamically self-managed "window size" with values in $\mathbb{R}_+$ until it is done. A fictitious window size of -1 will be assigned to OFF users. The state space for a user is $\mathbb{R}_+ \cup \{-1\}$, with the trace of the usual metric. The cemetery state (with resurrections) brings new problems with respect to Graham and Robert [7], but the choice of -1 simplifies notations.

The way ON users utilise the nodes depends only on their classes, and is given in terms of an *allocation matrix*

$$A = (A_{jk}, 1 \leq j \leq J,\ 1 \leq k \leq K)\,, \qquad A_{jk} \in \mathbb{R}_+\,,$$

as follows: at node j, the instantaneous weighted throughput of a class k user in state $w \in \mathbb{R}_+ \cup \{-1\}$ is given by

$$A_{jk} w^+ := A_{jk} w \mathbb{1}_{\{w \geq 0\}}\,,$$

which vanishes for OFF users and is proportional to state (window size) for ON users.

Hence the load at node j is given by the total weighted throughput of all users, and when the n-th of class k is in state $w_{n,k}$, by

$$u_j = \sum_{k=1}^{K} \sum_{n=1}^{N_k} A_{jk} w_{n,k}^+\,, \qquad 1 \leq j \leq J\,. \tag{1}$$

The node load vector

$$u = (u_j, 1 \leq j \leq J)$$

is an important descriptor of the congestion of the network, and is indirectly sensed by the users through their losses. It is a linear function of the states of the ON users.

In a simple example, mapping classes to routes,

$$A_{jk} = 1 \text{ if node } j \text{ is used by a class } k \text{ user, and } A_{jk} = 0 \text{ otherwise.}$$

More general matrices allow to model, *e.g.*, different utilisations of a given node by different classes, or packets of a given user taking varied routes in the network.

2.2. **Connection initiation and termination, congestion control.** Transitions between ON and OFF states are given by functions $\lambda_k : \mathbb{R}_+^J \to \mathbb{R}_+$ and $\mu_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ and laws α_k on $\mathbb{R}_+$, for $1 \leq k \leq K$. When the node load vector is in state $u = (u_j, 1 \leq j \leq J)$, see (1), then in class k:

- any OFF user (in state -1) is turned ON at instantaneous rate $\lambda_k(u)$ with new state w chosen according to $\alpha_k(dw)$, and establishes a connection for data transmission,
- any ON user in state $w \geq 0$ is turned OFF (to state -1) at rate $\mu_k(w, u)$, and terminates the connection.

The throughput of connections are determined by the instantaneous states of the users that have initiated them; in a simple example, their initial laws α_k are delta masses, such as δ_0. They cause congestion at the nodes of the network, see (1), which is controlled by ON users by regulating their state (throughput). This involves functions $a_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$

and $b_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ and constants $r_k \in [0,1]$, for $1 \le k \le K$; when the node load vector is $u = (u_j, 1 \le j \le J)$ and the state of a class k user is $w \ge 0$, then instantaneously:

- the state increases continuously at speed $a_k(w,u)$,
- the state jumps from w to $r_k w$ at rate $b_k(w,u)$, corresponding to the occurrence of a loss in the transmission due to congestion (usually $r_k = 1/2$).

The study in Graham and Robert [7] considered permanent connections, corresponding to having $\mu_k(w,u) \equiv 0$ and all users ON at time 0. Then, users can be identified with connections and have state space $\mathbb{R}_+$, and there is no use for λ_k and α_k.

Natural special forms for the parameters. A simple case is when the throughput rate is proportional to the state (window size) of any user, so that there are functions ν_k and β_k from $\mathbb{R}_+^J$ to $\mathbb{R}_+$ such that

$$\mu_k(w,u) = w\nu_k(u)\,, \qquad b_k(w,u) = w\beta_k(u)\,. \tag{2}$$

A special case has constant functions $\lambda_k(u) \equiv \lambda_k > 0$ and $\nu_k(u) \equiv \nu_k > 0$ (with slight abuse of notation), so that any class k user goes ON at rate λ_k, is required to transmit a quantity of data which follows an exponential law with parameter ν_k at instantaneous throughput rate given by its window, and then goes OFF. Also, a natural case has

$$a_k(w,u) = a_k(u) \tag{3}$$

regardless of w (with slight abuse of notation), and related to the inverse of the round-trip time (RTT) in the network when its load is given by the vector u.

A natural sub-case of the above is when

$$a_k(u) = \left(\tau_k + \sum_{j=1}^{J} t_{jk}(u_j)\right)^{-1}, \qquad \beta_k(u) = \delta_k + \sum_{j=1}^{J} d_{jk}(u_j)\,, \tag{4}$$

where, for class k users, $\tau_k > 0$ is the RTT between source and destination and $\delta_k \ge 0$ is the loss rate in a non-congested network, and $t_{jk}(u_j) \ge 0$ is the supplementary RTT delay and $d_{jk}(u_j) \ge 0$ loss rate at node j when its load is u_j. We may expect $t_{jk}(u_j)$ and $d_{jk}(u_j)$ to behave linearly in u_j, at least for large u_j.

Hence, care will be taken to include cases where $b_k(w,u)$ may have a *quadratic* behavior in the assumptions to be made.

General forms of the parameters λ_k, μ_k, α_k, a_k, b_k, and r_k are used to model, *e.g.*, balking (at arrival and later) in a congested network, more complex relations between window size and throughput and losses, sequences of losses due to congestion, slow start, RED, etc. For instance, a last very simple case is $\lambda_k(u) \equiv \lambda_k > 0$ and $\mu_k(w,u) \equiv \mu_k > 0$, corresponding to someone popping in and out to browse the Internet in his spare time.

2.3. **Markov process and its SDE representation, mean-field scaling.**

A Markov process describing this model is given by

$$W^N(t) = (W^N_{n,k}(t), 1 \le n \le N_k,\, 1 \le k \le K)\,, \qquad t \ge 0\,,$$

where

$$W^N_{n,k}(t) \in \mathbb{R}_+ \cup \{-1\}$$

is the state of the n-th user of class k at time t. We choose to represent it as the solution of the Itô-Skorohod stochastic differential equation (SDE)

$$
\left\{
\begin{aligned}
&dW^N_{n,k}(t) = \mathbb{1}_{\{W^N_{n,k}(t-)=-1\}} \int (1+w)\mathbb{1}_{\{0<z<\lambda_k(U^N(t-))\}}\, \mathcal{A}_{n,k}(dw,dz,dt) \\
&\qquad\qquad + \mathbb{1}_{\{W^N_{n,k}(t-)\geq 0\}} \Big[a_k(W^N_{n,k}(t-), U^N(t-))\, dt \\
&\qquad\qquad\quad - (1-r_k)W^N_{n,k}(t-) \int \mathbb{1}_{\{0<z<b_k(W^N_{n,k}(t-),U^N(t-))\}}\, \mathcal{N}_{n,k}(dz,dt) \\
&\qquad\qquad\quad - (1+W^N_{n,k}(t-)) \int \mathbb{1}_{\{0<z<\mu_k(W^N_{n,k}(t-),U^N(t-))\}}\, \mathcal{D}_{n,k}(dz,dt) \Big] , \\
&1 \leq k \leq K\,,\ 1 \leq n \leq N_k\,; \\
&U^N(t) = (U^N_j(t), 1 \leq j \leq J)\,, \qquad U^N_j(t) = \sum_{k=1}^{K} A_{jk} \sum_{i=1}^{N_k} W^N_{n,k}(t)^+\,,
\end{aligned}
\right.
\tag{5}
$$

driven by Poisson point processes $\mathcal{A}_{n,k}$ on $\mathbb{R}^3_+$ with intensity measure $\alpha_k(dw)\,dz\,dt$ and $\mathcal{N}_{n,k}$ and $\mathcal{D}_{n,k}$ on $\mathbb{R}^2_+$ with intensity measure $dz\,dt$, all independent. See Graham and Robert [7] for a discussion of the related simpler model for permanent connections. The following can be proved with standard arguments.

Proposition 2.1. *If the functions a_k are Lipschitz and the functions b_k, λ_k, and μ_k are locally bounded, $1 \leq k \leq K$, then there is pathwise existence and uniqueness of solution for the stochastic differential equation* (5)*, and the corresponding Markov process is well defined.*

2.4. **The mean-field asymptotic regime.** The high-dimensional system of coupled stochastic differential equations (5) does not seem to be mathematically tractable as such, and an asymptotic study, reducing its dimension $|N|$ to the number of classes K, is used to investigate the qualitative properties of solutions.

We consider the mean-field scaling in which

$$
N_k \to \infty\,, \quad \frac{N_k}{|N|} := \frac{N_k}{N_1 + \cdots + N_K} \to p_k\,, \qquad 1 \leq k \leq K\,,
\tag{6}
$$

(necessarily $(p_k)_{1\leq k\leq K}$ is a probability vector), and moreover the capacities of the resources are scaled by a factor $|N|$, *i.e.*, U^N is replaced by $\overline{U}^N = \frac{1}{|N|}U^N$ in (5) so as to have a non-trivial limit. This procedure leads to the *rescaled* SDE

$$
(7)\qquad
\begin{cases}
dW_{n,k}^N(t) = \mathbb{1}_{\{W_{n,k}^N(t-)=-1\}} \displaystyle\int (1+w)\mathbb{1}_{\{0<z<\lambda_k(\overline{U}^N(t-))\}}\, \mathcal{A}_{n,k}(dw,dz,dt) \\
\qquad\qquad + \mathbb{1}_{\{W_{n,k}^N(t-)\geq 0\}} \Big[a_k\big(W_{n,k}^N(t-),\overline{U}^N(t-)\big)\, dt \\
\qquad\qquad\quad - (1-r_k)W_{n,k}^N(t-) \displaystyle\int \mathbb{1}_{\{0<z<b_k(W_{n,k}^N(t-),\overline{U}^N(t-))\}}\, \mathcal{N}_{n,k}(dz,dt) \\
\qquad\qquad\quad - (1+W_{n,k}^N(t-)) \displaystyle\int \mathbb{1}_{\{0<z<\mu_k(W_{n,k}^N(t-),\overline{U}^N(t-))\}}\, \mathcal{D}_{n,k}(dz,dt) \Big]\,, \\
1\leq k\leq K\,,\ 1\leq n\leq N_k\,; \\
\overline{U}^N(t) = \big(\overline{U}_j^N(t), 1\leq j\leq J\big)\,, \qquad \overline{U}_j^N(t) = \dfrac{1}{|N|}U_j^N(t) = \displaystyle\sum_{k=1}^K A_{jk}\frac{N_k}{|N|}\overline{W}_k^N(t)\,, \\
\overline{W}_k^N(t) = \dfrac{1}{N_k}\displaystyle\sum_{n=1}^{N_k} W_{n,k}^N(t)^+\,.
\end{cases}
$$

This is a multi-class mean-field system, in interaction through the vector

$$\overline{W}^N(t) = \big(\overline{W}_k^N(t), 1\leq k\leq K\big)$$

of the empirical means of the class throughputs, *via* the vector $\overline{U}^N(t)$ of the *scaled loads*. For $1\leq k\leq K$, it is natural to introduce the *empirical measure* for the processes associated to class k users, given by

$$\Lambda_k^N = \frac{1}{N_k}\sum_{n=1}^{N_k}\delta_{(W_{n,k}^N(t),t\geq 0)}$$

where $\delta_{(x(t),t\geq 0)}$ denotes the Dirac mass at the function $t\in\mathbb{R}_+\mapsto x(t)\in\mathbb{R}_+\cup\{-1\}$, and in particular $\overline{W}_k^N(t) = \langle w^+,\Lambda_k^N(t)(dw)\rangle$, where $\Lambda_k^N(t)$ is the marginal of Λ_k^N at time t.

2.5. **Notations and conventions.** For $x=(x_m, 1\leq m\leq M)\in\mathbb{R}^M$ and $v:\mathbb{R}_+\to\mathbb{R}^M$ for some $M\in\mathbb{N}$ and $T>0$, let

$$\|x\| = \max_{1\leq m\leq M}|x_m|\,, \qquad \|v\|_T = \sup_{0\leq s\leq T}\|v(s)\|\,.$$

Depending on the context, the function $v=(v(t),t\geq 0)$ may be denoted

$$(v_m(t), 1\leq m\leq M, t\geq 0) \ \text{ or } \ (v_m(t), 1\leq m\leq M) \ \text{ or } \ (v(t))\,.$$

If H is a complete metric space, $\mathcal{D}(\mathbb{R}_+,H)$ denotes the Skorohod space of functions with values in H, right-continuous with left limits at any point of $\mathbb{R}_+$, see Billingsley [3].

The r.v. Λ_k^N has values in the set $\mathcal{P}(\mathcal{D}(\mathbb{R}_+,\mathbb{R}_+\cup\{-1\}))$ of probability measures on $\mathcal{D}(\mathbb{R}_+,\mathbb{R}_+\cup\{-1\})$.

3. Analysis of the nonlinear limit process

3.1. **Heuristic derivation of the mean-field limit.** For (7), the symmetry properties within each user class lead us to expect a mean-field convergence phenomenon as N gets large, for appropriately converging initial conditions:

- the processes $W_{n,k}^N$ for $1\leq k\leq K$ and $1\leq n\leq N_k$ *should* become independent and converge in law to processes W_k (depending only on the class k),

- and the empirical measure Λ_k^N *should* converge in law to a deterministic limit given by the law of the same process W_k,

where the stochastic process

$$(W(t), t \geq 0) = ((W_k(t), t \geq 0), 1 \leq k \leq K)$$

is the solution of the nonlinear, or McKean-Vlasov, Itô-Skorohod SDE

$$\begin{cases} dW_k(t) = \mathbb{1}_{\{W_k(t-)=-1\}} \displaystyle\int (1+w) \mathbb{1}_{\{0<z<\lambda_k(u_W(t))\}} \mathcal{A}_k(dw, dz, dt) \\ \qquad\qquad + \mathbb{1}_{\{W_k(t-)\geq 0\}} \Bigg[a_k\left(W_k(t-), u_W(t)\right) dt \\ \qquad\qquad\qquad - (1-r_k) W_k(t-) \displaystyle\int \mathbb{1}_{\{0<z<b_k(W_k(t-), u_W(t))\}} \mathcal{N}_k(dz, dt) \\ \qquad\qquad\qquad - (1+W_k(t-)) \displaystyle\int \mathbb{1}_{\{0<z<\mu_k(W_k(t-), u_W(t))\}} \mathcal{D}_k(dz, dt) \Bigg], \\ 1 \leq k \leq K\,; \\ u_W(t) = (u_{W,j}(t), 1 \leq j \leq J)\,, \qquad u_{W,j}(t) = \displaystyle\sum_{k=1}^K A_{jk} p_k \mathbb{E}(W_k(t)^+)\,, \end{cases} \tag{8}$$

driven by Poisson point processes $\mathcal{A}_k$ on $\mathbb{R}_+^3$ with intensity measure $\alpha_k(dw)\, dz\, dt$ and $\mathcal{N}_k$ and $\mathcal{D}_k$ on $\mathbb{R}_+^2$ with intensity measure $dz\, dt$, all independent.

The interaction between coordinates depends on the mean throughput vector

$$(u_W(t), t \geq 0)$$

which is a linear functional of the mean class output vector

$$\mathbb{E}(W(t)^+) = \mathbb{E}(W_k(t)^+, 1 \leq k \leq K) = \langle w^+, \mathcal{L}(W(t)) \rangle\,.$$

In particular, the infinitesimal generator of the process $(W(t), t \geq 0)$ depends, at time t, on the *law* of $W(t)$ and not only on the value taken by the sample path, as it is usually the case. Using this generator, there is a nonlinear martingale problem formulation for the weak interpretation of Equation (8).

Some properties of the solutions of the SDE (8) are analyzed in this section. The mean-field convergence results will be the topic of Section 4.

3.2. **Existence and uniqueness results.** The following proposition is the central technical result used to establish the existence and uniqueness of solutions of Equation (8), as well as the mean-field convergence result. For this last purpose, it is convenient (but not essential) to obtain a bound $C(t)$ which does not depend on $\|u'\|_t$.

Proposition 3.1. *For $1 \leq k \leq K$, let the functions a_k be bounded and a_k, b_k, λ_k, and μ_k be Lipschitz, and the laws α_k have a first moment $m_k = \int w\, \alpha_k(dw) < \infty$, and let $\mathcal{A}_k$ be Poisson point processes on $\mathbb{R}_+^3$ with intensity measure $\alpha_k(dw)\, dz\, dt$ and $\mathcal{N}_k$ and $\mathcal{D}_k$ be Poisson point processes on $\mathbb{R}_+^2$ with intensity measure $dz\, dt$, all forming an independent family. For any $u = (u(t), t \geq 0)$ in $C(\mathbb{R}_+, \mathbb{R}^J)$ and $(\mathbb{R}_+ \cup \{-1\})^K$-valued random variable X_0 independent of the Poisson point processes, let*

$$\phi(X_0, u) = (X(t), t \geq 0) = ((X_k(t), t \geq 0), 1 \leq k \leq K)$$

be the solution starting at $X(0) = X_0$ *of the stochastic differential equation*

$$
(\mathcal{E}_u) \qquad \begin{cases} dX_k(t) = \mathbb{1}_{\{X_k(t-)=-1\}} \displaystyle\int (1+w) \mathbb{1}_{\{0<z<\lambda_k(u(t))\}} \, \mathcal{A}_k(dw, dz, dt) \\ \qquad + \mathbb{1}_{\{X_k(t-)\geq 0\}} \Big[a_k\left(X_k(t-), u(t)\right) \, dt \\ \qquad\qquad - (1-r_k) X_k(t-) \displaystyle\int \mathbb{1}_{\{0<z<b_k(X_k(t-),u(t))\}} \, \mathcal{N}_k(dz, dt) \\ \qquad\qquad - (1+X_k(t-)) \displaystyle\int \mathbb{1}_{\{0<z<\mu_k(X_k(t-),u(t))\}} \, \mathcal{D}_k(dz, dt) \Big] , \\ 1 \leq k \leq K \, . \end{cases}
$$

Let $u = (u(t), t \geq 0)$ *and* $u' = (u'(t), t \geq 0)$ *be in* $C(\mathbb{R}_+, \mathbb{R}^J)$*, and initial values* X_0 *and* X_0' *be independent of the Poisson point processes. Then, for all* $t \geq 0$*,*

$$
\mathbb{E}\left[\|\phi(X_0, u) - \phi(X_0', u')\|_t \,\Big|\, X_0, X_0' \right] \leq \left[\|X_0 - X_0'\| + \int_0^t C(s) \|u - u'\|_s \, ds \right] e^{tC(t)} \tag{9}
$$

where $C(t) = A(1 + t + \|u\|_t + \|X_0\| + \|X_0'\|)$ *for some constant* $A < \infty$ *depending only on* a_k*,* b_k*,* λ_k*,* μ_k*, and* m_k*,* $1 \leq k \leq K$*.*

Proof. Any solution $(Y(t), t \geq 0) = (Y_k(t), 1 \leq k \leq K, t \geq 0)$ of $(\mathcal{E}_u)$, for any u and initial condition, satisfies the *a priori* affine growth bound

$$
\mathbb{E}^0\left[Y_k(t)^+ \right] \leq \max(Y_k(0)^+, m_k) + \|a_k\| t \, , \tag{10}
$$

which notably allows to prove, classically under the present assumptions, existence and uniqueness of solutions for $(\mathcal{E}_u)$ recursively from jump instant to next jump instant, since it implies that these cannot accumulate. Let

$$
X(t) = \phi(X_0, u)(t) \, , \quad X'(t) = \phi(X_0', u')(t) \, , \qquad t \geq 0 \, ,
$$

and $\mathbb{E}^0$ denote the conditional expectation given X_0 and X_0'. Then

$$
\|X_k - X_k'\|_t \leq |X_k(0) - X_k'(0)| + I_1(t) + I_2(t) + I_3(t) + I_4(t) \tag{11}
$$

for

$$
\begin{aligned}
I_1(t) &= \int_0^t \int (1+w) \Big| \mathbb{1}_{\{X_k(s-)=-1\}} \mathbb{1}_{\{0<z<\lambda_k(u(s))\}} \\
&\qquad\qquad - \mathbb{1}_{\{X_k'(s-)=-1\}} \mathbb{1}_{\{0<z<\lambda_k(u'(s))\}} \Big| \, \mathcal{A}_k(dw, dz, ds) \, , \\
I_2(t) &= \int_0^t \Big| \mathbb{1}_{\{X_k(s-)\geq 0\}} a_k(X_k(s), u(s)) - \mathbb{1}_{\{X_k'(s-)\geq 0\}} a_k(X_k'(s), u'(s)) \Big| \, ds \, , \\
I_3(t) &= \int_0^t \int \Big| X_k(s-)^+ \mathbb{1}_{\{0<z<b_k(X_k(s-),u(s))\}} \\
&\qquad\qquad - X_k'(s-)^+ \mathbb{1}_{\{0<z<b_k(X_k'(s-),u'(s))\}} \Big| \, \mathcal{N}_k(dz, ds) \, , \\
I_4(t) &= \int_0^t \int \Big| (1 + X_k(s-)) \mathbb{1}_{\{0<z<\mu_k(X_k(s-),u(s))\}} \\
&\qquad\qquad - (1 + X_k'(s-)) \mathbb{1}_{\{0<z<\mu_k(X_k'(s-),u'(s))\}} \Big| \, \mathcal{D}_k(dz, ds) \, ,
\end{aligned}
$$

where the fact that $\mathbb{1}_{\{x\geq 0\}} x = x^+$ and $\mathbb{1}_{\{x \geq 0\}}(1+x) = 1+x$ for x in $\mathbb{R}_+ \cup \{-1\}$ is used for regularizing some integrands.

Compensating the Poisson point process $\mathcal{A}_k$ and the inequality

$$\left|\mathbb{1}_{\{x=-1\}} - \mathbb{1}_{\{x'=-1\}}\right| = \left|\mathbb{1}_{\{x\geq 0\}} - \mathbb{1}_{\{x'\geq 0\}}\right| \leq |x-x'| \,, \quad x,x' \in \mathbb{R}_+ \cup \{-1\}\,, \tag{12}$$

(the functions $\mathbb{1}_{\mathbb{R}_+}$ and $\mathbb{1}_{\{-1\}}$ are 1-Lipschitz on $\mathbb{R}_+ \cup \{-1\}$) yield

$$\begin{aligned}
\mathbb{E}^0[I_1(t)] &= \mathbb{E}^0\biggl[\int_0^t \int (1+w)\bigl|\mathbb{1}_{\{X_k(s)=-1\}}\mathbb{1}_{\{0<z<\lambda_k(u(s))\}} \\
&\qquad\qquad - \mathbb{1}_{\{X'_k(s)=-1\}}\mathbb{1}_{\{0<z<\lambda_k(u'(s))\}}\bigr|\,\alpha_k(dw)\,dz\,ds\biggr] \\
&\leq (1+m_k)\mathbb{E}^0\biggl[\int_0^t \int \Bigl[\bigl|\mathbb{1}_{\{X_k(s)=-1\}} - \mathbb{1}_{\{X'_k(s)=-1\}}\bigr|\mathbb{1}_{\{0<z<\lambda_k(u(s))\}} \\
&\qquad\qquad + \mathbb{1}_{\{X'_k(s)=-1\}}\bigl|\mathbb{1}_{\{0<z<\lambda_k(u(s))\}} - \mathbb{1}_{\{0<z<\lambda_k(u'(s))\}}\bigr|\Bigr]\,dz\,ds\biggr] \\
&\leq (1+m_k)\mathbb{E}^0\biggl[\int_0^t \int \Bigl[\bigl|X_k(s) - X'_k(s)\bigr|\lambda_k(u(s)) + \bigl|\lambda_k(u(s)) - \lambda_k(u'(s))\bigr|\Bigr]\,ds\biggr]\,.
\end{aligned}$$

The singularities due to the cemetery state -1 now cause further difficulties. A *key step* in the proof is that, since the functions a_k, b_k, and μ_k, defined on $\mathbb{R}_+ \times \mathbb{R}_+^J$, can be extended *arbitrarily* to $(\mathbb{R}_+ \cup \{-1\}) \times \mathbb{R}_+^J$ without altering I_2, I_3, and I_4, we may introduce adequate Lipschitz extensions for such functions f by setting $f(-1,u) = f(0,u)$, *i.e.*,

$$f(x,u) = f(x^+,u)\,, \qquad (x,u) \in (\mathbb{R}_+ \cup \{-1\}) \times \mathbb{R}_+^J\,.$$

Since $x \mapsto x^+$ is 1-Lipschitz, the extensions are Lipschitz on $(\mathbb{R}_+ \cup \{-1\}) \times \mathbb{R}_+^J$ with same Lipschitz coefficients as the original function on $\mathbb{R}_+ \times \mathbb{R}_+^J$. (Note that the extension vanishing on $\{-1\} \times \mathbb{R}^J$ may well not be Lipschitz on $(\mathbb{R}_+ \cup \{-1\}) \times \mathbb{R}_+^J$, see for instance $(x,u) \in \mathbb{R}_+ \times \mathbb{R}^J \mapsto x+u$.)

In the sequel, we use these adequate Lipschitz extensions of a_k, b_k, and μ_k, which play an important role in regularizing the integrands, and invisibly simplify greatly the computations. Simple computations and (12) yield

$$\begin{aligned}
\mathbb{E}^0[I_2(t)] &= \mathbb{E}^0\biggl[\int_0^t \bigl|\mathbb{1}_{\{X_k(s)\geq 0\}}a_k(X_k(s),u(s)) - \mathbb{1}_{\{X'_k(s)\geq 0\}}a_k(X'_k(s),u'(s))\bigr|\,ds\biggr] \\
&\leq \mathbb{E}^0\biggl[\int_0^t \int \Bigl[\bigl|\mathbb{1}_{\{X_k(s)\geq 0\}} - \mathbb{1}_{\{X'_k(s)\geq 0\}}\bigr|a_k(X_k(s),u(s)) \\
&\qquad\qquad + \mathbb{1}_{\{X'_k(s)\geq 0\}}\bigl|a_k(X_k(s),u(s)) - a_k(X'_k(s),u'(s))\bigr|\Bigr]\,dz\,ds\biggr] \\
&\leq \mathbb{E}^0\biggl[\int_0^t \int \Bigl[\bigl|X_k(s) - X'_k(s)\bigr|a_k(X_k(s),u(s)) \\
&\qquad\qquad + \bigl|a_k(X_k(s),u(s)) - a_k(X'_k(s),u'(s))\bigr|\Bigr]\,dz\,ds\biggr]\,,
\end{aligned}$$

compensating the Poisson point process $\mathcal{N}_k$ yields

$$\begin{aligned}\mathbb{E}^0[I_3(t)] &= \mathbb{E}^0\left[\int_0^t \int \left|X_k(s)^+ \mathbb{1}_{\{0<z<b_k(X_k(s),u(s))\}} - X_k'(s)^+ \mathbb{1}_{\{0<z<b_k(X_k'(s),u'(s))\}}\right| dz\, ds\right] \\ &\leq \mathbb{E}^0\left[\int_0^t \int \Big[\left|X_k(s)^+ - X_k'(s)^+\right| \mathbb{1}_{\{0<z<b_k(X_k(s),u(s))\}}\right. \\ &\qquad\qquad \left. + X_k'(s)^+ \left|\mathbb{1}_{\{0<z<b_k(X_k(s),u(s))\}} - \mathbb{1}_{\{0<z<b_k(X_k'(s),u'(s))\}}\right|\Big] dz\, ds\right] \\ &\leq \mathbb{E}^0\left[\int_0^t \Big[\left|X_k(s) - X_k'(s)\right| b_k(X_k(s), u(s))\right. \\ &\qquad\qquad \left. + X_k'(s)^+ \left|b_k(X_k(s), u(s)) - b_k(X_k'(s), u'(s))\right|\Big] ds\right]\end{aligned}$$

and similarly, replacing x^+ by $1 + x$ in the computations,

$$\begin{aligned}\mathbb{E}^0[I_4(t)] \leq \mathbb{E}^0\Bigg[\int_0^t \Big[&\left|X_k(s) - X_k'(s)\right| \mu_k(X_k(s), u(s)) \\ &+ (1 + X_k'(s))\left|\mu_k(X_k(s), u(s)) - \mu_k(X_k'(s), u'(s))\right|\Big] ds\Bigg].\end{aligned}$$

Going back to (11), the bounds on $\mathbb{E}^0[I_1(t)]$, $\mathbb{E}^0[I_2(t)]$, $\mathbb{E}^0[I_3(t)]$, and $\mathbb{E}^0[I_4(t)]$, the affine growth bound (10), and the Lipschitz bounds which remain valid for the *extensions* of a_k, b_k and μ_k on $(\mathbb{R}_+ \cup \{-1\}) \times \mathbb{R}_+^J)$, yield

$$\mathbb{E}^0\left[\|X_k - X_k'\|_t\right] \leq |X_k(0) - X_k'(0)| + \int_0^t C_k(s) \left(\|u(s) - u'(s)\| + \mathbb{E}^0\left[\|X_k - X_k'\|_s\right]\right) ds$$

where $C_k(s) = A_k(1 + s + \|u\|_s + \|X_0\| + \|X_0'\|)$ for some constant $A_k < \infty$ depending only on a_k, b_k, λ_k, μ_k, and m_k, $1 \leq k \leq K$. The Gronwall Lemma then yields

$$\begin{aligned}\mathbb{E}^0\left[\|X_k - X_k'\|_t\right] &\leq \left[|X_k(0) - X_k'(0)| + \int_0^t C_k(s)\|u(s) - u'(s)\|\, ds\right] e^{\int_0^t C_k(s)\, ds} \\ &\leq \left[|X_k(0) - X_k'(0)| + \int_0^t C_k(s)\|u - u'\|_s\, ds\right] e^{tC_k(t)}\end{aligned}$$

and the proof of the proposition follows. □

One of the main motivations of this study is to obtain results which are valid for functions μ_k and b_k of the form given in (2) and (4). Such functions may have quadratic behavior, and Proposition 3.1 must be adapted to this case, which replaces $\|X_0\| + \|X_0'\|$ by $\|X_0\|^2 + \|X_0'\|^2$ in the bound for $C(t)$, but for simplicity this will be stated only when needed.

In order to control the evolution of $(W(t), t \geq 0)$ and describe its stationary behavior, the simple assumption that initial conditions be uniformly bounded is not satisfactory. For these purposes, exponential and Gaussian moment assumptions are introduced.

Condition (C) *It is said to hold for a family $\{X_0^\theta, \theta \in \Theta\}$ of $(\mathbb{R}_+ \cup \{-1\})^K$-valued r.v., for the functions $b_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ and $\mu_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ for $1 \leq k \leq K$, and for $\varepsilon > 0$, when at least one of the two following conditions is satisfied:*

(1) *The functions b_k and μ_k are Lipschitz for all $1 \leq k \leq K$, and the X_0^θ for $\theta \in \Theta$ have a uniform exponential moment of order ε:*

$$\sup_{\theta\in\Theta} \mathbb{E}\left[\exp(\varepsilon\|X_0^\theta\|)\right] < \infty\,.$$

(2) *There are Lipschitz functions* $\beta_k : \mathbb{R}_+^J \to \mathbb{R}_+$ *and* $\nu_k : \mathbb{R}_+^J \to \mathbb{R}_+$ *such that*

$$b_k(w,u) = w\beta_k(u)\,, \quad \mu_k(w,u) = w\nu_k(u)\,, \qquad 1 \le k \le K\,,$$

and the X_0^θ, $\theta \in \Theta$, *have a uniform Gaussian moment of order* ε*:*

$$\sup_{\theta\in\Theta} \mathbb{E}\big[\exp\big(\varepsilon \|X_0^\theta\|^2\big)\big] < \infty\,.$$

The following theorem establishes the existence and uniqueness result for Equation (8).

Theorem 3.2. *If the functions* $a_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ *are bounded and Lipschitz and* $\lambda_k : \mathbb{R}_+^J \to \mathbb{R}_+$ *are Lipschitz,* $1 \le k \le K$*, and if Condition (C) holds for the* $(\mathbb{R}_+ \cup \{-1\})^K$*-valued r.v.* W_0*, the functions* b_k *and* μ_k *for* $1 \le k \le K$*, and* $\varepsilon > 0$*, then there is pathwise existence and uniqueness of the solution* $(W(t), t \ge 0)$ *of the nonlinear stochastic differential equation* (8) *starting at* W_0*.*

In this case, the solution depends continuously on the initial condition in the following way: if $(W(t), t \ge 0)$ *and* $(W'(t), t \ge 0)$ *are solutions of* (8) *with respective initial conditions* W_0 *and* W_0' *having the same moment in Condition (C), then for* $T \ge 0$ *there exists a constant* A_T *such that*

$$\begin{aligned}(13)\quad \mathbb{E}\left[\|W - W'\|_T\right] &:= \mathbb{E}\left[\sup_{s\le T} \|W(s) - W'(s)\|\right] \\ &\le A_T \mathbb{E}\left[\|W_0 - W_0'\| e^{\varepsilon(\|W_0\|^\ell + \|W_0'\|^\ell)/2)}\right] e^{\mathbb{E}[\exp(\varepsilon(\|W_0\|^\ell + \|W_0'\|^\ell))]}\end{aligned}$$

where $\ell = 1$ *or* 2 *depending on whether the moment in Condition (C) is exponential or Gaussian.*

Proof. The proof follows *mutatis mutandis* the proof of Graham-Robert [7, Theorem 4.2], substituting Proposition 3.1 to [7, Proposition 4.1], and we give some details about this. If (1) of Condition (C) holds, then (9) (given in Proposition 3.1) is substituted to [7, (4.3)] (given in [7, Proposition 4.1]). If (2) of Condition (C) holds, then Proposition (3.1) and its proof must be adapted as in the end of the proof of [7, Theorem 4.2]: the term $\mathbb{E}^0[I_3(t)]$ must be bounded using

$$\begin{aligned}\big|b_k(X_k(s), u(s)) - b_k(X_k'(s), u'(s))\big| &= \big|X_k(s)\beta_k(u(s)) - X_k'(s)\beta_k(u'(s))\big| \\ &\le X_k(s)|\beta_k(u(s)) - \beta_k(u'(s))| + \beta_k(u'(s))|X_k(s) - X_k'(s)|\end{aligned}$$

and similarly for $\mathbb{E}^0[I_4(t)]$, and the rest of the proof is analogous, the supplementary multiplications by $X_k(s)$ and $X_k'(s)$ being handled by using the Gaussian moment assumption. □

4. Mean-field limit theorem for converging initial data

It is said that multi-indices $N = (N_k, 1 \le k \le K)$ go to infinity, denoted by $N \to \infty$, when

$$\min_{1\le k\le K} N_k \to \infty\,.$$

This is the case in the mean-field scaling, where (6) gives the relative growth rate of the numbers of class k users. In this context, the notions of exchangeability and chaoticity play a fundamental role. These properties are classical for single-class systems, see Aldous [1] and Sznitman [12] for example, but need to be extended to the present multi-class network.

Definition 1. *The family of r.v.* $(X_{n,k}, 1 \le n \le N_k, 1 \le k \le K)$ *is* multi-exchangeable *if its law is invariant under permutation of the indexes* within *the classes: for any permutations* σ_k *of* $\{1, \dots, N_k\}$ *for* $1 \le k \le K$*, there holds the equality of laws*

$$\mathcal{L}(X_{\sigma_k(n),k}, 1 \le n \le N_k, 1 \le k \le K) = \mathcal{L}(X_{n,k}, 1 \le n \le N_k, 1 \le k \le K)\,.$$

A family $\left(X^N_{n,k}, 1 \le n \le N_k, 1 \le k \le K\right)$ *of multi-class random variables indexed by* $N = (N_k, 1 \le k \le K) \in \mathbb{N}^K$ *is* $P_1 \otimes \cdots \otimes P_K$-multi-chaotic*, where each* P_k *is a probability measure, if for any* $m \ge 1$ *there holds the weak convergence of laws*

$$\lim_{N\to\infty} \mathcal{L}\left(X^N_{n,k}, 1 \le n \le m, 1 \le k \le K\right) = P_1^{\otimes m} \otimes \cdots \otimes P_K^{\otimes m}\,.$$

Hence, such a family of systems is multi-chaotic if and only if it becomes asymptotically independent with particles of class k having law P_k. A surprising result, not used in this paper, is that a family of multi-exchangeable multi-class systems is multi-chaotic *if and only if* the *restriction* to each class is chaotic, see Graham [6, Theorem 3].

The following theorem is the main mean-field convergence result. The underlying topology on the corresponding functional space is uniform convergence on compact sets.

Theorem 4.1. *It is assumed that:*

(1) *the mean-field regime* (6) *holds: the multi-indices* $N = (N_1, \dots, N_K)$ *go to infinity so that*
$$\lim_{N\to+\infty} \frac{N_k}{N_1 + \cdots + N_K} = p_k\,, \qquad 1 \le k \le K\,,$$
(2) *the* $\mathbb{R}_+ \cup \{-1\}$*-valued random variables*
$$\left(W^N_{n,k}(0), 1 \le n \le N_k, 1 \le k \le K\right)$$
are multi-exchangeable and $P_{1,0} \otimes \cdots \otimes P_{K,0}$*-multi-chaotic, where* $P_{k,0}$ *is a probability distribution on* $\mathbb{R}_+ \cup \{-1\}$ *for* $1 \le k \le K$*,*

(3) *the functions* $a_k : \mathbb{R}_+ \times \mathbb{R}_+^J \to \mathbb{R}_+$ *are bounded and Lipschitz, the functions* $\lambda_k : \mathbb{R}_+^J \to \mathbb{R}_+$ *are Lipschitz, and Condition (C) holds for the random variables*
$$\{W^N_1(0), N \in \mathbb{N}^K\} = \{(W^N_{1,k}(0), 1 \le k \le K), N \in \mathbb{N}^K\}\,,$$
the functions b_k *and* μ_k *for* $1 \le k \le K$*, and* $\varepsilon > 0$*.*

Then, for $1 \le k \le K$*, in the sense of processes*

$$\lim_{N\to+\infty} \mathbb{E}\left| \frac{1}{N_k} \sum_{n=1}^{N_k} W^N_{n,k}(t) - \mathbb{E}(W_k(t)) \right| = 0$$

and the family of processes

$$\left((W^N_{n,k}(t), t \ge 0), 1 \le n \le N_k, 1 \le k \le K\right)$$

given by the solutions of the SDE (7) *starting at the initial conditions* $\left(W^N_{n,k}(0)\right)$ *is multi-exchangeable and* P_W*-multi-chaotic, where* $P_W = P_{W_1} \otimes \cdots \otimes P_{W_K}$ *is the law of the solution*

$$(W(t), t \ge 0) = ((W_k(t), t \ge 0), 1 \le k \le K)$$

of the nonlinear SDE (8) *with initial distribution* $P_{1,0} \otimes \cdots \otimes P_{K,0}$*.*

In particular, $(W^N_{n,\cdot}(t), t \ge 0) := ((W^N_{n,k}(t), 1 \le k \le K), t \ge 0)$ *converges in distribution to* $(W(t), t \ge 0) = ((W_k(t), 1 \le k \le K), t \ge 0)$ *when* $N \to \infty$*, all* $n \ge 1$*.*

Proof. As in the proof for Theorem 3.2, the proof follows the proof of Graham-Robert [7, Theorem 5.1], substituting Proposition 3.1 to [7, Proposition 4.1] if (1) of Condition (C) holds, and adapting appropriately Proposition 3.1 if (2) of Condition (C) holds. □

The following is a simple consequence of the above result and of exchangeability.

Corollary 4.2. *Under the assumptions and notations of Theorem 4.1, the convergence in law of the empirical distributions*

$$\lim_{N\to\infty} \Lambda_k^N = P_{W_k}\,, \qquad 1 \le k \le K\,,$$

holds for the weak topology on $\mathcal{P}(\mathcal{D}(\mathbb{R}_+, \mathbb{R}_+ \cup \{-1\}))$ *with* $\mathcal{D}(\mathbb{R}_+, \mathbb{R}_+ \cup \{-1\})$ *endowed with the Skorohod topology.*

5. Equilibrium behavior

5.1. **Stationary distributions for the nonlinear SDE.** The basic equilibrium properties of $(W(t), t \ge 0) = ((W_k(t), t \ge 0), 1 \le k \le K)$, solution of the nonlinear SDE (8) describing the limit user evolution, are now succinctly studied, in particular its stationary distributions.

Note that, due to the non-linearity, $(W(t), t \ge 0)$ is *not* a homogeneous Markov process, and therefore the classical theory concerning the convergence toward equilibrium of Markov processes does not apply.

The coordinates of the process $(W(t), t \ge 0)$ evolve independently, and are coupled through the presence in the coefficients of Equation (8) of the load vector $(u_W(t), t \ge 0)$ given by

$$u_W(t) = (u_{W,j}(t), 1 \le j \le J)\,, \qquad u_{W,j}(t) = \sum_{k=1}^{K} A_{jk} p_k \mathbb{E}(W_k(t)^+)\,. \tag{1}$$

If a stationary distribution exists, and is taken as the initial distribution, then $u_W(t)$ is constant in t, and hence $(W(t), t \ge 0)$ is a stationary process with independent coordinates.

One begins with the analysis of the equilibrium of a single process $(W_k(t), t \ge 0))$ when the load vector $u_W(t)$ is replaced by a constant u, then examines the fixed-point equation resulting from the coupling of the coordinates through Relation (1), which in equilibrium does not depend on t.

5.2. **Invariant measures for a generic process.** Throughout this section a load vector $u = (u_j, 1 \le j \le J) \in \mathbb{R}_+^J$ is fixed without further mention, and simplified notations are used, dropping the index k from notations. Normal notation will be resumed in the following section.

One studies the invariant measures for the classic Itô-Skorohod SDE on $\mathbb{R}_+ \cup \{-1\}$

$$\begin{aligned} dW(t) &= \mathbb{1}_{\{W(t-)=-1\}} \int (1+w) \mathbb{1}_{\{0<z<\lambda(u)\}}\, \mathcal{A}(dw, dz, dt) \\ &\quad + \mathbb{1}_{\{W(t-)\ge 0\}} \Big[a\left(W(t-), u\right)\, dt - (1-r) W(t-) \int \mathbb{1}_{\{0<z<b(W(t-),u)\}}\, \mathcal{N}(dz, dt) \\ &\qquad - (1 + W(t-)) \int \mathbb{1}_{\{0<z<\mu(W(t-),u)\}}\, \mathcal{D}(dz, dt) \Big] \end{aligned} \tag{2}$$

driven by Poisson point processes $\mathcal{A}$ on $\mathbb{R}_+^3$ with intensity measure $\alpha(dw) dz dt$, and $\mathcal{D}$ and $\mathcal{N}$ on $\mathbb{R}_+^2$ with intensity measure $dzdt$. An important element of the study will be the SDE on $\mathbb{R}_+$, corresponding to the evolution of a permanent connection,

$$dV(t) = a\left(V(t-), u\right)\, dt - (1-r) V(t-) \int \mathbb{1}_{\{0<z<b(V(t-),u)\}}\, \mathcal{N}(dz, dt)\,. \tag{3}$$

For definitions and results on Harris ergodicity, see Nummelin [10], Asmussen [2]. Actually, classical positive recurrent state techniques are used.

Theorem 5.1. *Assume that*

(1) *the function* $x \mapsto a(x,u)$ *is Lipschitz bounded, and* $\inf\{a(x,u) : x \geq 0\} > 0$,
(2) *the function* $x \mapsto b(x,u)$ *is locally bounded, and* $\inf\{b(x,u) : x \geq x_0\} > 0$ *for some* $x_0 \geq 0$.

Then, for given initial conditions, there exists a unique solution $(W(t), t \geq 0)$ *of* (2)*, and a unique solution* $(V(t), t \geq 0)$ *of* (3)*, and this defines Markov processes. Moreover,* $(V(t), t \geq 0)$ *has a positive recurrent state and hence a stationary distribution* π^V*, and is in particular Harris ergodic.*

Proof. The existence and uniqueness result of solution and the Markov property are classical. The hitting time of x_0 by $(V(t), t \geq 0)$ is defined by

$$T_{x_0} = \inf\{t > 0 : V(t) = x_0, \exists s < t, V(s) \neq x_0\}.$$

Let $\eta := \inf\{b(x,u) : x \geq x_0\}$, and $(\widetilde{V}(t), t \geq 0)$ be the solution of the SDE

$$d\widetilde{V}(t) = a(\widetilde{V}(t-), u)\, dt - (1-r)\widetilde{V}(t-) \int \mathbb{1}_{\{0<z<\eta\}}\, \mathcal{N}(dz, dt) \tag{4}$$

(existence and uniqueness are classical). If $\widetilde{V}(0) \geq V(0) \geq x_0$ then it is a simple matter to construct a coupling of $(\widetilde{V}(t), t \geq 0)$ and $(V(t), t \geq 0)$ such that the relation $\widetilde{V}(t) \geq V(t)$ holds as long as $V(t) \geq x_0$. The Markov process $(\widetilde{V}(t))$ is Harris ergodic; see Dumas *et al.* [5] for the existence of a stationary distribution satisfying the coupling property, and Chafaï *et al.* [4] for more on long-time convergence. The hitting time below x_0 is therefore integrable for $(\widetilde{V}(t))$, and hence for $(V(t))$.

Now if $V(0) < x_0$ then, since there are no upward jumps and $x \mapsto a(x,u)$ is lower bounded away from 0 and $x \mapsto b(x,u)$ is bounded on the compact set $[0, x_0]$, it is a simple matter to show that the hitting time of x_0 by $(V(t))$ is integrable. One therefore concludes that $\mathbb{E}_{x_0}(T_{x_0}) < \infty$, and x_0 is positive recurrent, and classical results conclude the proof. □

Note that $(W(t), t \geq 0))$ (as defined in this section) is an homogeneous Markov process, and can be seen as a solution of the SDE (3) killed at some random instants and regenerated after an exponentially distributed duration of time with parameter $\lambda(u)$. Because of this regenerative structure, the stationary distribution of the SDE (3) will not come into play as it did in Graham and Robert [7].

Theorem 5.2. *Let the functions* a *and* b *satisfy the assumptions of Theorem 5.1, and moreover* $\lambda(u) > 0$ *and* $\int \mu(x,u)\pi^V(dx) > 0$*. Let* $(V(t), t \geq 0))$ *be the solution of* (3) *starting at* $V(0)$ *with law* $\alpha(dw)$*, and* γ *be the positive measure on* $\mathbb{R}_+ \cup \{-1\}$ *defined by*

$$\int f(x)\, \gamma(dx) = f(-1) + \lambda(u) \int_0^{+\infty} \mathbb{E}\left(f(V(t)) \exp\left(-\int_0^t \mu(V(s), u)\, ds \right)\right) dt \tag{5}$$

for all non-negative Borel functions f *on* $\mathbb{R}_+ \cup \{-1\}$*. Then* γ *is the unique invariant measure of the Markov process* $(W(t), t \geq 0))$ *defined by* (2)*: for all* $t \geq 0$,

$$\int \mathbb{E}_x(f(W(t)))\gamma(dx) = \int f(x)\gamma(dx)$$

for all non-negative Borel functions f *on* $\mathbb{R}_+ \cup \{-1\}$.

Proof. Let the stopping time

$$\tau_{-1} = \inf\{t > 0 : W(t) = -1, \exists s < t, W(s) \neq -1\}$$

be the cycle time of $(W(t), t \geq 0)$ between two visits to -1. Classically, if $\mathbb{P}(\tau_{-1} = \infty) = 0$ then it is well known that the measure $\hat{\gamma}$ defined by

$$\int f(x)\,\hat{\gamma}(dx) = \mathbb{E}_{-1}\left(\int_0^{\tau_{-1}} f(W(s))\,ds\right) \tag{6}$$

for all non-negative Borel functions f on $\mathbb{R}_+ \cup \{-1\}$ is the unique invariant measure for $(W(t), t \geq 0)$, see *e.g.* Asmussen [2, Proposition 3.2], Robert [11, Proposition 8.12].

In the following, $W(0) = -1$. Then $(W(t), t \geq 0)$ remains in -1 for an exponentially distributed duration $E_{\lambda(u)}$ with parameter $\lambda(u)$, after which it jumps to an α-distributed state in $\mathbb{R}_+$. Then, its excursion in $\mathbb{R}_+$ has same distribution as $(V(t), t \geq 0)$ until it returns to -1. Thus, $\hat{\gamma}(\{-1\}) = 1/\lambda(u)$, and $\tau_{-1} = E_{\lambda(u)} + \tau$ where τ is the first instant of an inhomogeneous Poisson process on $\mathbb{R}_+$ with rate function $(\mu(V(t), u), t \geq 0)$, and

$$\mathbb{P}(\tau > t \mid V(s), 0 \leq s \leq t) = \exp\left(-\int_0^t \mu(V(s), u)\,ds\right).$$

Since $(V(t), t \geq 0)$ is Harris ergodic and $\int \mu(x, u)\pi^V(dx) > 0$, the corresponding ergodic theorem yields

$$\lim_{t \to +\infty} \int_0^t \mu(V(s), u)\,ds = +\infty\,, \text{ a.s.},$$

so that $\mathbb{P}(\tau = \infty) = 0$, and thus $\mathbb{P}(\tau_{-1} = \infty) = 0$ implying that $\hat{\gamma}$ is indeed the unique invariant measure. From the above,

$$\begin{aligned}
\mathbb{E}\left(\int_0^\tau f(V(t))\,dt\right) &= \mathbb{E}\left(\int_0^{+\infty} f(V(t))\mathbb{1}_{\{\tau > t\}}\,dt\right) \\
&= \mathbb{E}\left(\int_0^{+\infty} f(V(t))\mathbb{P}(\tau > t \,|\, V(s), 0 \leq s \leq t)\,dt\right) \\
&= \int_0^{+\infty} \mathbb{E}\left(f(V(t)) \exp\left(-\int_0^t \mu(V(s), u)\,ds\right)\right) dt
\end{aligned}$$

so that γ defined by (5) is equal to $\lambda(u)\hat{\gamma}$. ☐

If the measure γ has a finite mass then there is a unique stationary distribution given by $\pi := \gamma/(1 + \gamma(\mathbb{R}_+))$, else there is none.

Remark 1. *The measure γ is finite on all Borel functions f such that, for some $C > 0$, $|f(x)| \leq C\mu(x, u)$ for any $x \geq 0$. Indeed, since $t \to V(t)$ is continuous almost everywhere for the Lebesgue measure on $\mathbb{R}_+$, then*

$$\begin{aligned}
&\int_0^{+\infty} |f(V(t))| \exp\left(-\int_0^t \mu(V(s), u)\,ds\right) dt \\
&\qquad\qquad \leq C \int_0^{+\infty} \mu(V(t), u) \exp\left(-\int_0^t \mu(V(s), u)\,ds\right) dt \leq C.
\end{aligned}$$

In particular, if $\inf_{x \geq 0} \mu(x, u) > 0$ then γ has a finite mass.

5.3. **Fixed-point equations.** In this section, the study of the case of K classes of users is resumed, and regular notation is again used. It is assumed that the functions $a_k(\cdot, u)$ and $b_k(\cdot, u)$ satisfy the assumptions of Theorem 5.1, and that the process $(W_k(t), 1 \leq k \leq K)$ has a stationary distribution π on $\mathbb{R}^K$, and the corresponding load vector is denoted by $u = (u_j, 1 \leq j \leq J)$.

Because of the independence of the coordinates, π can be written as $\pi = \bigotimes_{k=1}^{K} \pi_{k,u}$ where $\pi_{k,u}$ is the stationary distribution of the solution $(W_k(t), t \geq 0)$ of SDE (2) with coefficients $\lambda_k(u)$ and r_k and functions $a_k(\cdot, u)$, $b_k(\cdot, u)$ and $\mu_k(\cdot, u)$. By Theorem 5.2,

$$
(7) \qquad \begin{cases} \displaystyle\int f(x)\, \pi_{k,u}(dx) = \frac{1}{1 + \lambda_k(u) Z_k(u)} f(-1) \\ \displaystyle\quad + \frac{\lambda_k(u)}{1 + \lambda_k(u) Z_k(u)} \int_0^{+\infty} \mathbb{E}\left(f(V_{k,u}(t)) \exp\left(- \int_0^t \mu_k(V_{k,u}(s), u)\, ds \right) \right) dt\,, \\ \displaystyle\quad \text{with } Z_k(u) = \int_0^{+\infty} \mathbb{E}\left(\exp\left(- \int_0^t \mu_k(V_{k,u}(s), u)\, ds \right) \right) dt\,, \end{cases}
$$

for all non-negative Borel functions f on $\mathbb{R} \cup \{-1\}$, where $V_{k,u}$ is the unique solution of the SDE (3) (with index k added to the functions) with initial distribution $\alpha_k(dw)$, and Z_k is the appropriate normalization constant. Hence, by rewriting (1) one gets the following theorem.

Theorem 5.3. *Let the functions a_k and b_k and λ_k and μ_k satisfy the assumptions of Theorem 5.2 for any $1 \leq k \leq K$ and $u \in \mathbb{R}_+^J$. Then, any stationary distribution of the nonlinear SDE* (8) *can be written as $\pi = \bigotimes_{k=1}^{K} \pi_{k,u}$, where $\pi_{k,u}$ is defined by Equation* (7) *and $u = (u_j, 1 \leq j \leq J) \in \mathbb{R}_+^J$ is a solution of the fixed point equation*

$$
(8) \qquad u_j = \sum_{k=1}^{K} A_{jk} p_k \frac{\lambda_k(u)}{1 + \lambda_k(u) Z_k(u)} \int_0^{+\infty} \mathbb{E}\left(V_{k,u}(t) \exp\left(- \int_0^t \mu_k(V_{k,u}(s), u)\, ds \right) \right) dt
$$

where $V_{k,u}$ is the solution of the SDE (3) *associated to r_k and the functions $a_k(\cdot, u)$, $b_k(\cdot, u)$, $\mu_k(\cdot, u)$ with initial distribution $\alpha_k(dw)$ and $Z_k(u)$ is the corresponding normalizing constant given in Equation* (7).

The characterization of the invariant distributions involves the transitory behavior of the solution $(V_{k,u}(t), t \geq 0)$, which leads to a more complex and less explicit expression for the fixed-point equation in comparison to the case of permanent connections investigated in Graham and Robert [7].

One concludes with two examples for the service rates (μ_k) under the assumptions of the above theorem.

Constant service rate. It is assumed that $\mu_k(x, u) = \mu_k(u)$ for $x \geq 0$ and $u \in \mathbb{R}_+^J$. Equation (8) is in this case

$$
u_j = \sum_{k=1}^{K} A_{jk} p_k \frac{\lambda_k(u)}{1 + \lambda_k(u)/\mu_k(u)} \int_0^{+\infty} \mathbb{E}\left[V_{k,u}(t) \right] e^{-\mu_k(u) t}\, dt, \quad 1 \leq j \leq J.
$$

The integral in the right hand side of the above equation is related to the resolvent of the Markov process $(V_{k,u}(t))$.

Linear service rate. It is assumed that $\mu_k(x, u) = x \mu_k(u)$ for $x \geq 0$ and $u \in \mathbb{R}_+^J$. Equation (8) becomes

$$
u_j = \sum_{k=1}^{K} A_{jk} p_k \frac{\lambda_k(u)/\mu_k(u)}{1 + \lambda_k(u) Z_k(u)},
$$

with

$$
Z_k(u) = \int_0^{+\infty} \mathbb{E}\left(\exp\left(-\mu_k(u) \int_0^t V_{k,u}(s)\, ds \right) \right) dt.
$$

Future Work. As it may be seen from Relation (8), the explicit representation of the invariant distribution involves the distribution of the solution of the SDE (3) associated to a permanent connection. If the equilibrium behavior of this process is fully understood, its transient characteristics are not well known. See Chafaï *et al.* [4] for the rate of convergence to equilibrium. The case of constant rate shows that one should have an expression of the resolvent of this process. This is, in our view, an interesting challenging problem of this domain.

References

[1] David Aldous, *Exchangeability and related topics*, École d'été de Probabilités de Saint-Flour XIII, Lecture Notes in Mathematics, vol. 1117, Springer Verlag, 1985, pp. 1–198.

[2] Søren Asmussen, *Applied probability and queues*, second ed., Applications of Mathematics (New York), vol. 51, Springer-Verlag, New York, 2003, Stochastic Modelling and Applied Probability.

[3] Patrick Billingsley, *Convergence of probability measures*, Wiley series in probability and mathematical statistics, John Wiley & Sons Ltd, New York, 1968.

[4] Djalil Chafaï, Florent Malrieu, and Katy Paroux, *On the long time behavior of the TCP window size process*, Stochastic Processes and their Applications **120** (2010), no. 8, 1518 – 1534.

[5] Vincent Dumas, Fabrice Guillemin, and Philippe Robert, *A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms*, Advances in Applied Probability **34** (2002), no. 1, 85–111.

[6] Carl Graham, *Chaoticity for multi-class systems and exchangeability within classes*, Journal of Applied Probability **45** (2008), 1196–1203.

[7] Carl Graham and Philippe Robert, *Interacting multi-class transmissions in large stochastic networks*, Annals of Applied Probability **19** (2009), no. 6, 2334–2361.

[8] Carl Graham, Philippe Robert, and Maaike Verloop, *Stability properties of networks with interacting tcp flows*, Network Control and Optimization, Proceedings NET-COOP 2009, 2009, pp. 1–15.

[9] Fabrice Guillemin, Philippe Robert, and Bert Zwart, *AIMD algorithms and exponential functionals*, Annals of Applied Probability **14** (2004), no. 1, 90–117.

[10] Esa Nummelin, *General irreducible Markov chains and nonnegative operators*, Cambridge University Press, Cambridge, 1984.

[11] Philippe Robert, *Stochastic networks and queues*, Stochastic Modelling and Applied Probability Series, vol. 52, Springer, New-York, June 2003.

[12] A.S. Sznitman, *Topics in propagation of chaos*, École d'été de Saint-Flour, Lecture Notes in Maths, vol. 1464, Springer Verlag, 1989, pp. 167–243.

(C. Graham) UMR 7641 CNRS — École Polytechnique, Route de Saclay, 91128 Palaiseau, France

E-mail address: carl@cmapx.polytechnique.fr

(Ph. Robert) INRIA Paris — Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay, France

E-mail address: Philippe.Robert@inria.fr

URL: http://www-rocq.inria.fr/~robert