Stochastics and Dynamics
Vol. 21, No. 3 (2021) 2140009 (23 pages)

DOI: 10.1142/S0219493721400098

World Scientific
www.worldscientific.com

Internet congestion control: From stochastic to dynamical models

José M. Amigó*, Angel Giménez†, Oscar Martínez-Bonastre‡ and José Valero§

Centro de Investigación Operativa,
Universidad Miguel Hernández de Elche,
Avda. de la Universidad s/n, 03202 Elche, Spain
**jm.amigo@umh.es*
†*a.gimenez@umh.es*
‡*oscar.martinez@umh.es*
§*jvalero@umh.es*

Received 22 January 2020
Revised 8 February 2021
Accepted 10 February 2021
Published 15 June 2021

This paper is dedicated to Manfred Denker on the occasion of his 75th birthday.

Since its inception, control of data congestion on the Internet has been based on stochastic models. One of the first such models was Random Early Detection. Later, this model was reformulated as a dynamical system, with the average queue sizes at a router's buffer being the states. Recently, the dynamical model has been generalized to improve global stability. In this paper we review the original stochastic model and both nonlinear models of Random Early Detection with a two-fold objective: (i) illustrate how a random model can be "smoothed out" to a deterministic one through data aggregation and (ii) how this translation can shed light into complex processes such as the Internet data traffic. Furthermore, this paper contains new materials concerning the occurrence of chaos, bifurcation diagrams, Lyapunov exponents and global stability robustness with respect to control parameters. The results reviewed and reported here are expected to help design an active queue management algorithm in real conditions, that is, when system parameters such as the number of users and the round-trip time of the data packets change over time. The topic also illustrates the much-needed synergy of a theoretical approach, practical intuition and numerical simulations in engineering.

Keywords: Internet congestion control; adaptive queue management; random early detection; discrete-time dynamical systems; global stability; robust setting of control parameters.

AMS Subject Classification: 37E05, 39A50, 37G35, 37N35

1. Introduction

Internet Protocols (IP) networks include memory buffers to manage the network traffic by implementing a queue algorithm into buffers. Obviously, the length of this queue is upper bounded by the size of the buffer, and they are not effective in their purpose of managing bursts of packets if they run at full capacity. Therefore, in a scenario of congestion some actions have to be taken to avoid overflow or a system collapse. An Active Queue Management (AQM) is an algorithm acting on the queue to control its length and, thus, achieve an efficient control of the system congestion. Such management also enables the transmission control protocols (TCPs) to share links properly.

The Random Early Detection (RED) algorithm [12] was the first formal and complete AQM system implemented in TCP/IP networks [2]. One of the purposes of designing the RED queue management algorithm was to detect the beginning of traffic congestion well in advance. This would allow the system to act in two ways: starting to gradually discard packets and reducing the transmission rate to that node, both actions aimed at avoiding scenarios of massive packet discarding, system collapse or overflow.

The congestion detection indicator used by the RED queue management is an Exponentially Weighted Moving Average (EWMA) of the queue length. EWMA actually acts as a low-pass filter that smooths the instantaneous queue length bursts. The degree of smoothing is determined by a weight w. The algorithm uses a decision mechanism based on a linear probability distribution to determine when packets are rejected. When a packet reaches the queue, the algorithm proceeds as follows. If the weighted average of the queue length is less than a minimum threshold $q_{\min}$, no action is taken and the packet simply remains in the queue. An average queue length between $q_{\min}$ and another higher threshold $q_{\max}$ is interpreted as congestion, in which case an early drop command is executed. More specifically, the RED algorithm drops an incoming packet with a probability that depends on the EWMA. This means that adequate $q_{\min}$ and $q_{\max}$ values are a requirement for good performance. Finally, if the average queue length is greater than the maximum threshold $q_{\max}$, then it is understood that the congestion is persistent, and packets are dropped to avoid a lasting full queue. In this case, the packets are dropped with probability 1.

RED has been the most studied AQM to date, and has been the basis for the development of new AQM systems. This is so not only because it was the first one developed in the Internet community, but also because of several drawbacks that this algorithm entails, some of which have not yet been successfully resolved. In particular, it is difficult to adjust the RED parameters for adequate performance due to their high sensitivity. Moreover, even if the RED parameters are adjusted properly, they are also very sensitive to the network conditions. Thus, a set of parameters can work perfectly for a given load of data traffic but not be suitable if the load varies slightly. Obviously, this is not a desirable feature since Internet traffic

conditions change rapidly. This is, by far, what has most affected a widespread implementation of RED. There are many proposals in the literature to overcome these difficulties via modifications of the RED algorithm, e.g., [4–8, 13].

Numerical experiments and simulations have usually been the main tool for setting parameters, also in the papers cited above. As a result, many of the conclusions published in the literature are based only on numerical results, thus lacking the necessary mathematical rigor to support them. The first theoretical study of RED we are aware of was done by Ranjan *et al.* [19]. This was possible because these authors reformulated the original, stochastic RED model as a discrete-time dynamical model, thus allowing them to apply the powerful methods of nonlinear dynamics. In particular, the authors showed that their dynamical RED model was chaotic. This first dynamical model was generalized in [3, 9, 10] with the objective of enhancing the controllability of the model. To this end, two new control parameters where introduced via the probability law for dropping incoming packets. Moreover, the analysis of the generalized model in [3] focused on global stability rather than local stability like in [19]. Numerical results (e.g., bifurcation diagrams and parametric domains of robustness) suggest that, as expected, the generalized model is more stable than the original one. This is a welcome feature for the prospective design of an AQM in real conditions, that is, when system parameters such as the number of users and the round-trip time change over time.

In this paper, we review the original RED dynamical model [19] and the main results of the generalized dynamical model [3] with a two-fold objective: (i) illustrate how a random model can be "smoothed out" to a deterministic one through data aggregation (sometimes with much ingenuity), and (ii) how this translation can provide useful insights into complex processes such as data traffic on the Internet. Furthermore, this paper contains new materials concerning the occurrence of chaos, bifurcation diagrams, Lyapunov exponents and global stability robustness with respect to control parameters, as we detail in short. The topic (Internet congestion control) illustrates also the much-needed synergy of a theoretical approach, practical intuition and numerical simulations in engineering.

The rest of this paper is organized as follows. In Sec. 2.1, we present the probabilistic RED algorithm, and in Sec. 2.2, how to derive a nonlinear model by trading off detailed information (queue size) for coarse-grained information (average queue length). Our presentation is based on [19]. In Sec. 3.1, we follow [3] to generalize the said nonlinear model with the purpose of increasing the number of control parameters and thus the control leverage. We have also added in Sec. 3.2 the proof that the generalized RED dynamics can be chaotic (depending on the parameters) in the sense of Li–Yorke. The main results in [3] concerning the global stability of the unique fixed point of the RED dynamic are summarized in Sec. 4, while Sec. 5 recaps a few more particular results with a potential to be implemented in real congestion control. The numerical simulations presented in Sec. 6, all of them new, deal principally with bifurcation diagrams and Lyapunov exponents (Sec. 6.1),

as well as with the robustness of the fixed point with respect to the new control parameters (Sec. 6.2). Section 7 contains the concluding remarks and an outlook.

Finally, for the convenience of the reader, we list the main symbols used in this paper.

System parameters

B Buffer size (maximum number of packets that the buffer of Router 1 can store).
K System constant usually set equal to $\sqrt{3/2}$.
C Capacity of the channel (the maximum amount of error-free information that can be transmitted over the channel per unit time).
N Number of users connected to a Router 1 which shares an Internet link (bottleneck link) with Router 2.
M Packet size.

System variables

q^{ave} Average queue length at Router 1.
p Probability of dropping an incoming packet.

Control parameters

$p_{\max}$ Maximum packet drop probability.
$q_{\max}$ Upper thresholds of q^{ave} for accepting and dropping an incoming packet.
$q_{\min}$ Lower thresholds of q^{ave} for accepting and dropping an incoming packet.
w Averaging weight for the update of q^{ave} $(0 < w < 1)$.
α New control parameter of the generalized RED model.
β New control parameter of the generalized RED model.

Dynamical core

θ_l Lower threshold of the dynamical core (the dynamics is affine below θ_l).
θ_r Upper threshold of the dynamical core (the dynamics is linear above θ_r).

2. Random Early Detection (RED)

We follow [19] for the description of RED and the derivation of its discrete-time dynamical model. Our main concern is the mathematical content of the model, hence technical details will be kept to a minimum.

2.1. *Probabilistic RED algorithm*

Consider the following communication network throughout. N users (i.e. N uniform TCP connections) are connected to a Router 1 which shares an Internet link (bottleneck link) with Router 2. The *capacity* of this channel (i.e. the maximum amount of error-free information that can be transmitted over the channel per unit time) is C. Further parameters of the system are the *packet size* M, the *round-trip*

time (propagation delay with no queuing delay) d of the packets, and the *buffer size* B of Router 1. So, B is the maximum number of packets that the buffer of Router 1 can store. The *queue length* or *size* q at the buffer at a given time is the number of packets in the buffer at that time, so $q = 0, 1, \ldots, B$.

In the RED model, the probability p of dropping an incoming packet at a router depends on the "average" queue length $q^{\mathtt{ave}}$ (to be defined below) as follows:

$$p(q^{\mathtt{ave}}) = \begin{cases} 0 & \text{if } q^{\mathtt{ave}} < q_{\min}, \\ 1 & \text{if } q^{\mathtt{ave}} > q_{\max}, \\ \dfrac{q^{\mathtt{ave}} - q_{\min}}{q_{\max} - q_{\min}} p_{\max} & \text{otherwise.} \end{cases} \tag{2.1}$$

Thus, $q_{\min}$ and $q_{\max}$ are the lower and upper thresholds of $q^{\mathtt{ave}}$ for accepting and dropping an incoming packet, respectively, and $p_{\max}$ is the selected drop probability when $q^{\mathtt{ave}} = q_{\max}$, i.e. the maximum packet drop probability. The average queue length is updated at the time of a packet arrival according to the averaging

$$q^{\mathtt{ave}}_{\mathtt{new}} = (1 - w) q^{\mathtt{ave}}_{\mathtt{old}} + w q^{\mathtt{cur}} \tag{2.2}$$

between the previous average queue length $q^{\mathtt{ave}}_{\mathtt{old}}$ and the current queue length $q^{\mathtt{cur}}$, where $0 < w < 1$ is the *averaging weight*. The initial average queue length is set in an unspecified way; from then on, the average queue length is defined by the rule (2.2). The higher w, the faster the RED mechanism reacts to the actual buffer occupancy. In practice w is taken rather small, typically $\lesssim 0.2$ [21].

The reader interested in probabilistic AQM models is referred to [2].

2.2. *Nonlinear RED model*

On the way to a discrete-time nonlinear dynamical model, let $q^{\mathtt{ave}}_n$ be the average queue length at time $n = 1, 2, \ldots,$ Then, the packet drop probability at time n, p_n, determines the queue length at time $n + 1$, q_{n+1}. In turn, q_{n+1} is used to compute $q^{\mathtt{ave}}_{n+1}$ according to the averaging rule (2.2). Finally, $q^{\mathtt{ave}}_{n+1}$ determines the drop probability at time $n + 1$, p_{n+1}, via Equation (2.1). In sum, the update of p_n has this step-wise structure:

$$\begin{cases} q_{n+1} = G(p_n), \\ q^{\mathtt{ave}}_{n+1} = (1 - w) q^{\mathtt{ave}}_n + w q_{n+1}, \\ p_{n+1} = H(q^{\mathtt{ave}}_{n+1}), \end{cases} \tag{2.3}$$

where $H(q^{\mathtt{ave}}_{n+1}) = p(q^{\mathtt{ave}}_{n+1})$ is given in (2.1). From (2.3) it follows:

$$q^{\mathtt{ave}}_{n+1} = (1 - w) q^{\mathtt{ave}}_n + w G(H(q^{\mathtt{ave}}_n)). \tag{2.4}$$

So this equation will define a dynamic model for the average queue length once the function G is determined.

It was shown in [14, 16–18] that, if the packet size M, the time delay d and the packet loss probability p are known, the stationary throughput of a typical TCP connection can be approximated by

$$T(p,d) = \frac{MK}{\sqrt{p}d} + o\left(\frac{1}{\sqrt{p}}\right), \tag{2.5}$$

where K is a constant usually set equal to $\sqrt{3/2}$. Therefore, the *steady-state packet drop probability* p_r such that the aggregate throughput of connections equals the link capacity, is given by

$$N\frac{MK}{\sqrt{p_r}d} = C. \tag{2.6}$$

This is the smallest drop probability such that $p_n \geq p_r$ entails $q_{n+1} = 0$. By (2.6),

$$p_r = \left(\frac{NMK}{Cd}\right)^2. \tag{2.7}$$

The corresponding average queue length q_r^{ave} such that $q_n^{\mathsf{ave}} > q_r^{\mathsf{ave}}$ implies $q_{n+1} = 0$, follows from (2.1):

$$q_r^{\mathsf{ave}} = \begin{cases} \dfrac{q_{\max} - q_{\min}}{p_{\max}} p_r + q_{\min} & \text{if } p_r \leq p_{\max}, \\ q_{\max} & \text{otherwise.} \end{cases} \tag{2.8}$$

Note that

$$q_r^{\mathsf{ave}} \leq q_{\max}. \tag{2.9}$$

Otherwise, if $p_n < p_r$ (i.e. $q_{n+1} > 0$) and $q_{n+1} \leq B$, the Internet link capacity between Routers 1 and 2 is fully utilized and the round-trip time of a packet arriving at time $n+1$ is augmented from d (no queuing delay) to $d + Mq_{n+1}/C$. From

$$N\frac{MK}{\sqrt{p_n}(d + \frac{Mq_{n+1}}{C})} = C \tag{2.10}$$

(see (2.6)), we obtain

$$q_{n+1} = \frac{C}{M}\left(\frac{NKM}{\sqrt{p_n}C} - d\right). \tag{2.11}$$

Since q_{n+1} is a strictly decreasing function of p_n, there exists $\min\{p_n, n \geq 1 : q_{n+1} = B\} =: p_l$, where

$$p_l = \left(\frac{NMK}{BM + Cd}\right)^2 \tag{2.12}$$

by (2.11). The corresponding average queue length is

$$q_l^{\mathsf{ave}} = \frac{q_{\max} - q_{\min}}{p_{\max}} p_l + q_{\min} = \frac{q_{\max} - q_{\min}}{p_{\max}} \left(\frac{NMK}{BM + Cd}\right)^2 + q_{\min}. \tag{2.13}$$

Therefore, $q_{n+1} = B$ for $p_n \leq p_l$, i.e. $q_n^{\mathtt{ave}} \leq q_l^{\mathtt{ave}}$. Note that

$$q_l^{\mathtt{ave}} > q_{\min}. \tag{2.14}$$

Comparison of (2.13) with (2.8) shows that

$$q_l^{\mathtt{ave}} < q_r^{\mathtt{ave}} \tag{2.15}$$

if $p_r < p_{\max}$. For $q_l^{\mathtt{ave}} < q_r^{\mathtt{ave}}$ to hold also when $p_r \geq p_{\max}$ (i.e. $q_r^{\mathtt{ave}} = q_{\max}$), it is necessary that

$$\left(\frac{NMK}{BM + Cd}\right)^2 < p_{\max}. \tag{2.16}$$

From Eq. (2.11) and the definition of p_r, Eq. (2.7), and p_l, Eq. (2.13), it follows that the function G in (2.3) is defined as

$$q_{n+1} = G(p_n) = \begin{cases} B & \text{if } p_n \leq p_l, \\ \dfrac{NK}{\sqrt{p_n}} - \dfrac{Cd}{M} & \text{if } p_l < p_n < p_r, \\ 0 & \text{if } p_n \geq p_r. \end{cases} \tag{2.17}$$

Finally, plugging (2.17) into (2.4) we conclude that the time evolution of the average queue length is given by the nonlinear dynamic

$$q_{n+1}^{\mathtt{ave}} = \begin{cases} (1-w)q_n^{\mathtt{ave}} + wB & \text{if } q_n^{\mathtt{ave}} \leq q_l^{\mathtt{ave}}, \\ (1-w)q_n^{\mathtt{ave}} + w\left(\dfrac{NK}{\sqrt{p_n}} - \dfrac{Cd}{M}\right) & \text{if } q_l^{\mathtt{ave}} < q_n^{\mathtt{ave}} < q_r^{\mathtt{ave}}, \\ (1-w)q_n^{\mathtt{ave}} & \text{if } q_n^{\mathtt{ave}} \geq q_r^{\mathtt{ave}}, \end{cases} \tag{2.18}$$

where see (2.1),

$$p_n = p(q_n^{\mathtt{ave}}) = \frac{q_n^{\mathtt{ave}} - q_{\min}}{q_{\max} - q_{\min}} p_{\max} \tag{2.19}$$

since $q_l^{\mathtt{ave}} > q_{\min}$ by (2.14) and $q_r^{\mathtt{ave}} \leq q_{\max}$ by (2.9). So, f is affine on $[0, \theta_l]$, linear on $[\theta_r, 1]$, and monotone or multimodal on (θ_l, θ_r).

The mapping $q_n^{\mathtt{ave}} \mapsto q_{n+1}^{\mathtt{ave}}$ of the interval $[0, B]$, Eq. (2.18), is affine on $[0, q_l^{\mathtt{ave}}]$, $\cup$-convex on $[q_l^{\mathtt{ave}}, q_r^{\mathtt{ave}}]$ and linear on $[q_r^{\mathtt{ave}}, B]$, where $[q_l^{\mathtt{ave}}, q_r^{\mathtt{ave}}]$ can be shown to be a trapping interval of the ensuing dynamic. It is, therefore, not surprising that this dynamic may be chaotic, depending on the parameter settings [19]. This is bad news for the sake of a stable control of the Internet data traffic, which is our main concern. The RED dynamical model (2.18) has 5 *system parameters* (N, C, d, M, B) and 4 *user parameters* $(p_{\max}, q_{\min}, q_{\max}, w)$. The former are constrained by condition (2.16). The latter are also called *control parameters* because they can be tuned at will to stabilize the dynamic if necessary. Among the system parameters, the most critical ones are N and d because they may change in real time. When it comes to the implementation of adaptive control, the averaging weight w is the most practical

choice. In the next section we will add to w two more control parameters to improve the stability properties of the RED dynamical model (2.18).

3. A Generalized RED Dynamical Model

In this section, we generalize the dynamic (2.18) by introducing two additional parameters in the probability law (2.19). We start by simplifying the notation.

3.1. *Mathematical formulation*

For a more compact notation, we introduce normalized state variables and thresholds,

$$x_n = \frac{q_n^{\mathtt{ave}}}{B}, \quad x_{\min} = \frac{q_{\min}}{B}, \quad x_{\max} = \frac{q_{\max}}{B}, \tag{3.1}$$

as well as the dimensionless constants

$$A_1 = \frac{NK}{\sqrt{p_{\max}}B}, \quad A_2 = \frac{Cd}{MB}. \tag{3.2}$$

Actually, out of the parameters in (3.2) only C, d and M are dimensional (given, say, in kilobytes per second, seconds and kilobytes, respectively); the other parameters are pure numbers. The constant K will be set equal to $\sqrt{3/2}$ for definiteness, see (2.5). According to Eq. (2.16), A_1 and A_2 are related as follows.

Proposition 3.1. *The constants A_1 and A_2 of the RED model* (2.18) *are subject to the constraint*

$$A_1 < A_2 + 1. \tag{3.3}$$

Inequality (3.3) is assumed throughout this paper. Since $p_{\max}$ appears in the denominator of A_1, the larger $p_{\max}$, the better for the constraint (3.3) to be fulfilled.

On the grounds explained in the Introduction, the RED dynamical model (2.18) was generalized in [3] by replacing the probability law (2.19) by

$$p(x_n) = I_{\alpha,\beta}(z(x_n)) \cdot p_{\max}, \tag{3.4}$$

where $I_{\alpha,\beta}(z)$, $0 \le z \le 1$, is the beta distribution function (or normalized incomplete beta function),

$$I_{\alpha,\beta}(z) = \frac{\mathfrak{B}(z;\alpha,\beta)}{\mathfrak{B}(1;\alpha,\beta)}, \quad \mathfrak{B}(z;\alpha,\beta) = \int_0^z t^{\alpha-1}(1-t)^{\beta-1}dt, \tag{3.5}$$

with $\alpha, \beta > 0$, and

$$z(x_n) = \frac{x_n - x_{\min}}{x_{\max} - x_{\min}}, \quad x_{\min} \le x_n \le x_{\max}. \tag{3.6}$$

Since $I_{1,1}(z) = z$, we recover the conventional RED model (2.18) for $\alpha = \beta = 1$. The beta distribution is related to the chi-square distribution [1]. By definition,

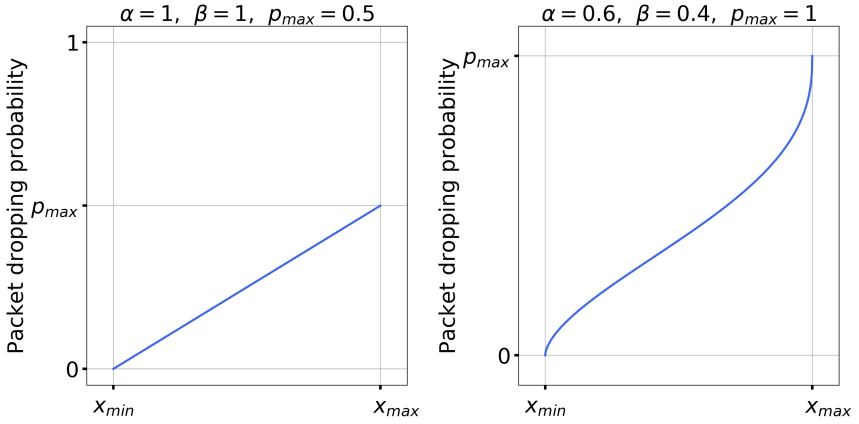

Fig. 1. Packet dropping probability, Eqs. (3.4)–(3.6), for the settings of the control parameters α, β and $p_{\max}$ displayed at the top of the panels.

$I_{\alpha,\beta}(z)$ is strictly increasing, hence invertible. Its inverse, $I^{-1}_{\alpha,\beta}(z)$, is also strictly increasing.

The packet dropping probability $p(x_n)$ is plotted in Figure 1 for different values of the control parameters α, β and $p_{\max}$. The left panel ($\alpha = \beta = 1$) corresponds to (2.19) with $p_{\max} = 0.5$.

Thus, we consider hereafter a dynamical system,

$$x_{n+1} = f(x_n), \tag{3.7}$$

where the mapping $f : [0,1] \to [0,1]$ is defined as

$$f(x) = \begin{cases} (1-w)x + w & \text{if } 0 \le x \le \theta_l, \\ (1-w)x + w\left(\dfrac{A_1}{\sqrt{I_{\alpha,\beta}(z(x))}} - A_2\right) & \text{if } \theta_l < x < \theta_r, \\ (1-w)x & \text{if } \theta_r \le q \le 1, \end{cases} \tag{3.8}$$

and the thresholds θ_l, θ_r are given by

$$\theta_l = (x_{\max} - x_{\min}) I^{-1}_{\alpha,\beta}(z_1) + x_{\min}, \quad z_1 = \left(\frac{A_1}{A_2 + 1}\right)^2 \tag{3.9}$$

($z_1 < 1$ by (2.16)) and

$$\theta_r = \begin{cases} (x_{\max} - x_{\min}) I^{-1}_{\alpha,\beta}(z_2) + x_{\min} & \text{if } z_2 = \left(\dfrac{A_1}{A_2}\right)^2 \le 1, \\ x_{\max} & \text{otherwise.} \end{cases} \tag{3.10}$$

Hence, $\theta_l > x_{\min}$,

$$\begin{cases} \theta_r < x_{\max} & \text{if } 0 < A_1 < A_2, \\ \theta_r = x_{\max} & \text{if } A_2 \le A_1 < A_2 + 1, \end{cases} \tag{3.11}$$

and $\theta_l < \theta_r$ by (3.3). In sum, f is affine on the interval $[0, \theta_l]$, linear on $[\theta_r, 1]$, and monotone or multimodal on (θ_l, θ_r).

The thresholds θ_l and θ_r are set so that f is continuous on $[0, 1]$, except when $A_1 > A_2$, in which case f is lower semicontinuous at θ_r. Indeed, if $A_1 > A_2$,

$$\begin{aligned} f(\theta_r-) = f(x_{\max}-) &= \lim_{x \to x_{\max}-} f(x) = (1-w)x_{\max} + w(A_1 - A_2) \\ &= f(x_{\max}) + w(A_1 - A_2) > f(x_{\max}) = f(\theta_r), \end{aligned} \tag{3.12}$$

where we used $I_{\alpha,\beta}(z(x_{\max})) = I_{\alpha,\beta}(1) = 1$, on the second line.

3.2. *Chaoticity of the generalized RED dynamics*

We will numerically show in Sec. 6 that the additional control parameters α and β of the generalized RED nonlinear model (3.8) do improve the stability of the dynamic. This is important to enable controllability also in a real environment, where the system parameters may change over time. It is worth noting that $p_{\max}$ and all five system parameters go into the dynamic through the constants A_1 and A_2, which amounts to a high degeneracy of the dynamic with respect to the system parameters. Note also that A_1, A_2 and the control parameters except w go into the thresholds θ_l and θ_r as well.

We prove next that the dynamical system defined by (3.8) can be chaotic, depending of the parameter setting. To be more specific, we are going to prove analytically that the dynamic $x_{n+1} = f(x_n)$ can exhibit chaos in the sense of Li–Yorke [15]. Numerical evidence in the form of bifurcation diagrams and Lyapunov exponents will be presented in Sec. 6.

We recall the main result of [15] first.

Theorem 3.1. *Let J be a closed interval and $F : J \to J$ a continuous mapping. Suppose that there is $x_0 \in J$ such that*

$$F^3(x_0) \le x_0 < F(x_0) < F^2(x_0) \tag{3.13}$$

or

$$F^3(x_0) \ge x_0 > F(x_0) > F^2(x_0). \tag{3.14}$$

Then

(1) *for every $k \in \mathbb{N}$ there is a periodic point in J with period k;*
(2) *there is an uncountable set $S \subset J$ containing no periodic points, which satisfies the following conditions:*
   (i) *For every $x, y \in S$, $x \neq y$,*

$$\limsup_{n \to \infty} |F^n(x) - F^n(y)| > 0$$

and

$$\liminf_{n \to \infty} |F^n(x) - F^n(y)| = 0.$$

(ii) *For every* $x \in S$ *and periodic point* $p \in J$,

$$\limsup_{n\to\infty} |F^n(x) - F^n(p)| = 0.$$

A dynamical system defined by a mapping $F : J \to J$ complying with the assumptions of Theorem 3.1 is said to be chaotic in the sense of Li–Yorke.

Theorem 3.2. *The generalized RED dynamic can be chaotic in the sense of Li–Yorke.*

Proof. We are going to apply Theorem 3.1. In our case, $J = [0, 1]$ and $F = f$, Eq. (3.8). For f to be continuous, we assume $A_1 \leq A_2$ for the time being.

Following [19], we choose $x_0 = \theta_r/(1-w)$ with $w < 1-\theta_r$ (so that $\theta_r < x_0 < 1$). Therefore,

$$f(x_0) = f\left(\frac{\theta_r}{1-w}\right) = \theta_r,$$

and

$$f^2(x_0) = f(\theta_r) = (1-w)\theta_r,$$

hence

$$x_0 > f(x_0) > f^2(x_0).$$

By (3.14) we only need that $f^3(x_0) \geq x_0$ to conclude the presence of chaos in the form stated in the points (1) and (2) of Theorem 3.1. There are two cases: $f^2(x_0) > \theta_l$, i.e. $w < (\theta_r - \theta_l)/\theta_r$ (Case I) and $f^2(x_0) \leq \theta_l$, i.e. $w \geq (\theta_r - \theta_l)/\theta_r$ (Case II).

In Case I, $\theta_l < f^2(x_0) < \theta_r$, hence $f^3(x_0) = f(f^2(x_0)) \geq x_0$ if and only if

$$f\left((1-w)\theta_r\right) = (1-w)^2\theta_r + w\left(\frac{A_1}{\sqrt{I_{\alpha,\beta}(z((1-w)\theta_r)}} - A_2\right) \geq \frac{\theta_r}{1-w}, \tag{3.15}$$

where $z((1-w)\theta_r) = ((1-w)\theta_r - x_{\min})/(x_{\max} - x_{\min})$. To verify that inequality (3.15) can be satisfied, use the following lower bound of the LHS of (3.15):

$$\begin{aligned} f\left((1-w)\theta_r\right) &\geq (1-w)^2\theta_r + \frac{wA_1}{\sqrt{I_{\alpha,\beta}(z((1-w)\theta_r)}} \\ &\geq (1-w)^2\theta_r + \frac{wA_1}{\sqrt{I_{\alpha,\beta}(z(\theta_r)}}, \end{aligned} \tag{3.16}$$

since $I_{\alpha,\beta}(z)^{-1/2}$ is a monotone decreasing function. Set now $w = \epsilon \ll 1$. Then, a sufficient condition for (3.15) to hold is that the lower bound (3.16) is greater than

$\frac{\theta_r}{1-\epsilon}$, that is,

$$(1-2\epsilon)\theta_r + \frac{\epsilon A_1}{\sqrt{I_{\alpha,\beta}(z(\theta_r)}} \geq (1+\epsilon)\theta_r$$

to first-order in ϵ. It follows that

$$\theta_r \sqrt{I_{\alpha,\beta}(z(\theta_r)} \leq \frac{A_1}{3} \tag{3.17}$$

suffices for $f^3(x_0) \geq x_0$ to hold, provided that w is sufficiently small.

In Case II, $f^2(x_0) \leq \theta_l$, hence $f^3(x_0) = f(f^2(x_0)) \geq x_0$ if and only if

$$f\left((1-w)\theta_r\right) = (1-w)^2\theta_r + w \geq \frac{\theta_r}{1-w},$$

hence

$$\theta_r \leq \frac{1-w}{3-w+w^2} < \frac{1}{3}. \tag{3.18}$$

This concludes the proof. □

In sum, compliance with (3.15) if $w < (\theta_r - \theta_l)/\theta_r$, or with (3.18) otherwise, implies that the generalized RED dynamics is chaotic in the sense of Li–Yorke. Unlike condition (3.15), which involves all the parameters of the model, condition (3.18) only involves the control parameter w. Actually, we expect the condition (3.17) to be the general criterion because, as mentioned in Sec. 2.1, typically $w \lesssim 0.2$. Needless to say, the possibility of a chaotic behavior poses a challenge to the control design.

4. Global Stability of the Generalized RED Model: Main Results

In this paper, we are interested in a stable and robust congestion control of the Internet data traffic based on the generalized RED model. This means that the system remains on, or evolves within the basin of attraction of, a fixed point for suitably chosen control parameters, also when the system parameters (grouped in the constants A_1 and A_2) and the control parameters (mainly w, α and β) change slightly. This way one can expect to keep the system in a neighborhood of the fixed point when both system parameters and (consequently) control parameters change in real time. Recall that $A_1 < A_2 + 1$ (Proposition 3.1).

To this end, this section summarizes the most important results on global stability of the dynamical system (3.8). Proofs and additional detail can be found in [3]. The study of robustness with respect to the control parameters α, β and w is deferred to Sec. 6.2.

Theorem 4.1. *The mapping f has a unique fixed point $x^* \in (\theta_l, \theta_r)$ if and only if $A_1 < A_2 + x_{\max}$; otherwise, it has none. Furthermore, x^* does not depend on w.*

To compute the fixed point x^* one has in general to solve numerically the equation $f(x) = x$ (with $\theta_l < x < \theta_r$ and $A_1 < A_2 + x_{\max}$), which leads to

$$\frac{A_1}{\sqrt{I_{\alpha,\beta}(z(x))}} = x + A_2, \tag{4.1}$$

where the LHS is a strictly decreasing function of x and the RHS is a strictly increasing function. It follows that x^* is an increasing function of A_1 (in particular, of N) by fixed A_2, whereas it is a decreasing function of A_2 (in particular, of d) by fixed A_1. In some case (e.g., $\alpha = \beta = 1$ and $\alpha = 2$, $\beta = 1$), Eq. (4.1) can be solved analytically.

Uniqueness of the fixed point x^* is good news for the control stability. But for the sake of controllability we need much more: global attractiveness of x^*.

Let $\mathcal{B}(S, f)$ denote the basin of attraction of a set $S \subset [0, 1]$, that is, $\mathcal{B}(S, f)$ consists of all points of $[0, 1]$ that asymptotically end up in S.

Theorem 4.2. *If (θ_l, θ_r) is invariant (i.e. $f((\theta_l, \theta_r)) \subset (\theta_l, \theta_r)$), then $\mathcal{B}((\theta_l, \theta_r), f) = [0, 1]$.*

Theorem 4.2 shows that the interval $[\theta_l, \theta_r]$ contains the non-wandering set for f when (θ_l, θ_r) is invariant, i.e. it is the dynamical core of the RED dynamic. Actually, the only way to prevent the orbit of $x_0 \in [0, \theta_l] \cup [\theta_r, 1]$ from getting trapped within (θ_l, θ_r) is that it is a preimage of an hypothetical periodic cycle $f(\theta_l) = \theta_r$ and $f(\theta_r) = \theta_l$. It is easy to show that $\theta_l + \theta_r \neq 1$ excludes that possibility.

If $\mathcal{B}(x^*, f) = [0, 1]$ we say that the fixed point x^* is a global attractor of f, what amounts to x^* being *globally stable*. Equivalently, we say also that f is globally stable. Let $f|_{(\theta_l,\theta_r)}$ denote as usual the restriction of the mapping f to the interval (θ_l, θ_r),

$$f|_{(\theta_l,\theta_r)}(x) = (1 - w)x + w\left(\frac{A_1}{\sqrt{I_{\alpha,\beta}(z(x))}} - A_2\right). \tag{4.2}$$

For further reference, note that $f|_{(\theta_l,\theta_r)}$ is a smooth function because $z(x) > 0$ for all $x \in (\theta_l, \theta_r)$ (since $x \geq \theta_l > x_{\min}$); in particular,

$$\begin{aligned} f'|_{(\theta_l,\theta_r)}(x) &= 1 - w\left(1 + \frac{A_1}{2(x_{\max} - x_{\min})}\frac{I'_{\alpha,\beta}(z(x))}{I_{\alpha,\beta}(z(x))^{3/2}}\right) \\ &< 1 - w, \end{aligned} \tag{4.3}$$

$0 < w < 1$, for all system and control parameters. If $A_1 > A_2$ (f discontinuous at θ_r) and $\beta < 1$, then $f'(\theta_r-) = f'(x_{\max}-) = -\infty$, so $f'|_{(\theta_l,\theta_r)}$ is upper bounded but not necessarily lower bounded.

Theorem 4.3. *Suppose $A_1 < A_2 + x_{\max}$ and let $x^* \in (\theta_l, \theta_r)$ be the unique fixed point of f. If*

(i) *(θ_l, θ_r) is invariant (so that $f|_{(\theta_l,\theta_r)}$ defines a dynamical system on (θ_l, θ_r)),*

(ii) $\theta_l + \theta_r \neq 1$ (*so that* $\{\theta_l, \theta_r\}$ *is not a periodic orbit*), *and*
(iii) $\mathcal{B}(x^*, f|_{(\theta_l,\theta_r)}) = (\theta_l, \theta_r)$ (i.e. x^* *is a global attractor of* $f|_{(\theta_l,\theta_r)}$), *then* x^* *is a global attractor of* f.

Therefore, to study the global stability of the RED dynamics it suffices to concentrate on $f|_{(\theta_l,\theta_r)}$, provided that conditions (i) and (ii) of Theorem 4.3 are satisfied. Practical considerations advice to limit the stability analysis to monotone and unimodal mappings over (θ_l, θ_r). A bimodal $f|_{(\theta_l,\theta_r)}$ can be seen in [3].

That a monotone $f|_{(\theta_l,\theta_r)}$ is strictly increasing or strictly decreasing depends on whether $f(\theta_l) < f(\theta_r-)$ or $f(\theta_l) > f(\theta_r-)$, respectively. This translates into the following condition on the averaging weight [3, Proposition 3]:

$$f(\theta_l) \gtrless f(\theta_r-) \Leftrightarrow w \gtrless \frac{\theta_r - \theta_l}{\theta_r - \theta_l + 1 - (A_1 - A_2)^+}. \tag{4.4}$$

As usual, the notation $(A_1 - A_2)^+$ stands for the positive part of $A_1 - A_2$ ($= \max\{A_1 - A_2, 0\}$), and similarly for other arguments.

Theorem 4.4. *If* $f|_{(\theta_l,\theta_r)}$ *is*

(a) *strictly increasing and* $A_1 < A_2 + x_{\max}$, *or*
(b) *strictly decreasing with* $f'|_{(\theta_l,\theta_r)}(x) > -1$ *and*

$$w \leq \min\left\{\frac{\theta_r - \theta_l}{1 - \theta_l}, \frac{\theta_r - \theta_l}{\theta_r - (A_1 - A_2)^+}\right\} \tag{4.5}$$

(*where* $\theta_r = x_{\max}$ *if* $(A_1 - A_2)^+ > 0$, *Eq.* (3.11)),

then $\mathcal{B}(q^*, f|_{(\theta_l,\theta_r)}) = (\theta_l, \theta_r)$.

The restriction $f'|_{(\theta_l,\theta_r)}(x) > -1$ in (b) can be replaced by the absence of 2-cycles, also at the endpoints $\{\theta_l, \theta_r\}$, should x^* be a global attractor. The latter can be done, as in Theorem 4.3, with the proviso $\theta_l + \theta_r \neq 1$. By Sharkovsky's theorem [20] applied to the continuous case ($A_1 \leq A_2$), if there are no periodic orbits of period 2, then there are no periodic orbits of any period.

Consider next the unimodal case. Specifically we suppose that $f|_{(\theta_l,\theta_r)}$ has a local extremum at $x_c \in (\theta_l, \theta_r)$.

Theorem 4.5. *Suppose that* (i) $w \leq \frac{\theta_r - \theta_l}{1 - \theta_l}$, (ii) $f|_{(\theta_l,\theta_r)}$ *has a local minimum at* $x_c \leq x^*$ (*so* $f'(x^*) \geq 0$), *and* (iii) $A_1 \leq A_2 + x_{\max}$. *Then* (θ_l, θ_r) *is invariant and* $\mathcal{B}(x^*, f|_{(\theta_l,\theta_r)}) = (\theta_l, \theta_r)$.

The case $x_c > x^*$ is more demanding.

Theorem 4.6. *The following holds*:

(a) *Suppose that* (i) $x^* > (A_1 - A_2)^+$, (ii) $f|_{(\theta_l,\theta_r)}$ *has a minimum at* $x_c > x^*$ (*so* $f'(x^*) < 0$), *and* (iii)

$$w \leq \min\left\{\frac{\theta_r - \theta_l}{1 - \theta_l}, \quad \frac{x^* - \theta_l}{x^* - (A_1 - A_2)^+}\right\}. \tag{4.6}$$

If $f'(x) > -1$ *for* $\theta_l < x < x_c$, *then* (θ_l, θ_r) *is invariant and* $\mathcal{B}(x^*, f|_{(\theta_l,\theta_r)}) = (\theta_l, \theta_r)$.

(b) *Part* (*a*) *holds also if* (*i*) *is replaced by* $x^* \leq (A_1 - A_2)^+ < x_{\max}$, *and* (*iii*) *is replaced by*

$$w \leq \frac{\theta_r - \theta_l}{1 - \theta_l}. \tag{4.7}$$

Similar results can be obtained when $f|_{(\theta_l,\theta_r)}$ has a local maximum. However, the assumptions in this case are more restrictive due to (4.3). Therefore, we will not consider this option hereafter. The same happens regarding the Schwarzian derivative

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2. \tag{4.8}$$

The condition $Sf|_{[\theta_l,\theta_r]}(x) < 0$ for all $x \neq x_c$, along with (i) $A_1 \leq A_2$ (so that f is continuous at θ_r), (ii) an invariant $[\theta_l, \theta_r]$, and (iii) $|f'(q^*)| \leq 1$, implies $\mathcal{B}(x^*, f|_{[\theta_l,\theta_r]}) = [\theta_l, \theta_r]$ [11, Proposition 1], but its implementation is quite involved and restrictive in parametric space.

5. Global Stability of the Generalized RED Model: Particular Results

When it comes to putting in place a control mechanism of the dynamics (3.8) that can be upgraded to an adaptive mechanism in real time, simplicity and computational speed is a must. This has been our guideline when selecting Theorems 4.4 to 4.6, from which one derives the sought global stability of x^*, i.e. $\mathcal{B}(x^*, f) = [0, 1]$, under the following three assumptions:

(i) (θ_l, θ_r) is invariant,
(ii) $\{\theta_l, \theta_r\}$ is not a periodic orbit, and
(iii) $f|_{(\theta_l,\theta_r)}$ has at most one minimum.

It turns out that a good compromise satisfying (i)–(iii) is to choose $f|_{(\theta_l,\theta_r)}$ $\cup$-convex. In particular, the convexity of f in the dynamical core (θ_l, θ_r) simplifies Theorems 4.4(b) and 4.6(a) as follows.

Theorem 5.1. *If* $f|_{(\theta_l,\theta_r)}$ *is* $\cup$*-convex, then the hypothesis* $f'|_{(\theta_l,\theta_r)}(x) > -1$ *in Theorem* 4.1(*b*) *may be replaced by* $f'(\theta_l+) \geq -1$.

Theorem 5.2. *If $f|_{(\theta_l,\theta_r)}$ is $\cup$-convex, along with $(A_1-A_2)^+ < x^*$ and $x_c > x^*$ (so $f'(x^*) < 0$), then the assumptions $f'|_{(\theta_l,q_c)}(x) > -1$ and (4.6) in Theorem 4.6(a) may be replaced by $f'(\theta_l+) \geq -1$ and*

$$w \leq \min\left\{\frac{\theta_r-\theta_l}{1-\theta_l}, \frac{x^*-\theta_l+\frac{1}{m}(\theta_l-(A_1-A_2)^+)^+}{x^*-(A_1-A_2)^+}\right\}, \tag{5.1}$$

respectively, where $m = (1-f'(x^*)/w > 1$.

Since $\frac{1}{m}(\theta_l-(A_1-A_2)^+)^+ \geq 0$, the bound (5.1) is indeed less restrictive than the bound (4.6). When $m \to \infty$ in (5.1) we recover (4.6), as it should.

General conditions for $f|_{(\theta_l,\theta_r)}$ to be $\cup$-convex (with one or no critical point) follow readily from the expression

$$f''|_{(\theta_l,\theta_r)}(x) = \frac{wA_1}{4(x_{\max}-x_{\min})}\frac{I'_{\alpha,\beta}(z(x))}{I_{\alpha,\beta}(z(x))^{3/2}}\left[3J_{\alpha,\beta}(z(x))-2h_{\alpha,\beta}(z(x))\right], \tag{5.2}$$

where $z(x) = (x-x_{\min})/(x_{\max}-x_{\min})$, and

$$J_{\alpha,\beta}(z) := \frac{I'_{\alpha,\beta}(z)}{I_{\alpha,\beta}(z)} > 0, \quad h_{\alpha,\beta}(z) := \frac{\alpha-1}{z}-\frac{\beta-1}{1-z}, \tag{5.3}$$

for all $z(\theta_l) < z < z(\theta_r)$.

Proposition 5.1. *Suppose that $3J_{\alpha,\beta}(z(x))-2h_{\alpha,\beta}(z(x)) > 0$ for $\theta_l < x < \theta_r$. It holds*:

(a) *$f_{(\theta_l,\theta_r)}$ is $\cup$-convex.*
(b) *If $f'(\theta_l+)\cdot f'(\theta_r-) > 0$, then $f|_{(\theta_l,\theta_r)}$ has no critical point; if $f'(\theta_l+)\cdot f'(\theta_r-) < 0$, then $f|_{(\theta_l,\theta_r)}$ has one critical point.*

There are a number of settings for the control parameters α and β that guarantee the sufficient condition $3J_{\alpha,\beta}(z(x))-2h_{\alpha,\beta}(z(x)) > 0$ for $f_{(\theta_l,\theta_r)}$ to be $\cup$-convex. The perhaps most useful ones are the following.

Proposition 5.2. *$f|_{(\theta_l,\theta_r)}$ is $\cup$-convex at x in the following cases*:

(a) *$\alpha \leq \beta$ and $0 < z(x) < \frac{\alpha}{\alpha+\beta}$.*
(b) *$\alpha > \beta$ and $0 < z(x) < \frac{\alpha+2}{\alpha+\beta+4}$.*

Therefore, the mapping $f|_{(\theta_l,\theta_r)}$ is $\cup$-convex whenever

$$z(\theta_r) = \frac{\theta_r-x_{\min}}{x_{\max}-x_{\min}} \leq \begin{cases}\dfrac{\alpha}{\alpha+\beta} & \text{if } \alpha \leq \beta,\\ \dfrac{\alpha+2}{\alpha+\beta+4} & \text{if } \alpha > \beta.\end{cases}$$

See [3] for other possibilities.

6. Numerical Simulations

The system parameters used as reference values in this section are

$$N = 1850, \quad C = 321,000 \text{ kBps}, \quad d = 0.012 \text{ s}, \\ M = 1 \text{ kB}, \quad B = 2000 \text{ packets}, \tag{6.1}$$

along with $K = \sqrt{3/2} \simeq 1.225$. Data (6.1) correspond to the Miguel Hernández University network and are the same as in [3]. The parameters A_1 and A_2 result then in

$$A_1 \simeq \frac{1.13}{\sqrt{p_{\max}}} \quad \text{and} \quad A_2 \simeq 1.93,$$

see (3.2), so the constraint $A_1 < A_2 + 1$, Proposition 3.1, holds for $p_{\max} \gtrsim 0.15$.

As for the control parameters: α, β and $p_{\max}$ will be specified in each figure, while the reference values for the remaining control parameters are

$$w = 0.15, \quad x_{\min} = 0.2, \quad x_{\max} = 0.6, \tag{6.2}$$

when kept fixed.

Next we present bifurcation diagrams of the RED dynamics and the corresponding Lyapunov exponents (Sec. 6.1), as well as a brief discussion of the robustness of the fixed point x^* with respect to the control parameters α, β (Sec. 6.2). The primary objective of Sec. 6.1 is to illustrate the effect of the control parameters α, β on two particular bifurcation diagrams. As for Sec. 6.2, robustness of the RED dynamics with respect to the control parameters in general, and with respect to α and β in particular, is essential to enable controllability under real conditions.

6.1. *Bifurcation diagrams and Lyapunov exponents*

The choice of appropriate parameters to ensure the stability of the dynamic (meaning that the fixed point x^* is a global attractor) is perhaps the biggest drawback of the RED algorithm. Therefore, it is of great interest to analyze the bifurcation diagram and the Lyapunov exponent for different parameters. Some of them were studied in [3], but the analysis of the two very important parameters $x_{\min}$ and $x_{\max}$ were omitted there for brevity. The mappings we use here are plotted in Fig. 2.

Figures 3 and 4 show the bifurcation diagram and the Lyapunov exponent with respect to $x_{\min}$ and $x_{\max}$, respectively, obtained with the values $\alpha = \beta = 1$ (upper row) and $\alpha = 0.6$, $\beta = 0.4$ (lower row) for the values $p_{\max} = 1$ (left column) and $p_{\max} = 0.5$ (right column). The settings of the other control parameters are given in the captions of the figures. For clarity, we also depict the endpoints θ_l, θ_r of the dynamical core, that encloses the fixed point x^*. The parametric grid used for the bifurcation diagrams has 2000 points; the orbits were 550 iterates long, the first 500 (the transient) having been discarded.

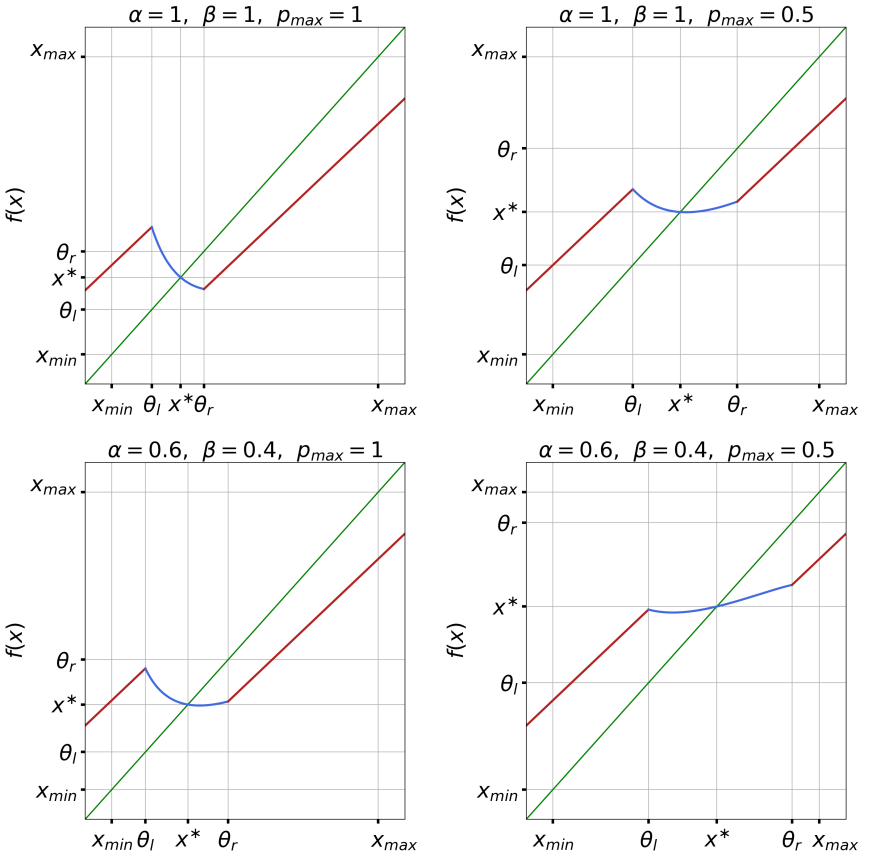

Fig. 2. Graphs of the mapping f, Eq. (3.8), for the values of α, β and $p_{\max}$ displayed at the top of the panels. Remaining parameters are given in (6.1) and (6.2). The restriction $f|_{(\theta_l,\theta_r)}$ is monotonically decreasing in the upper left panel, while it is $\cup$-convex in the other panels.

The most salient aspects of these figures can be summarized as follows:

(1) The bifurcations are direct for $x_{\min}$ (i.e. for growing values), whereas they are reverse for $x_{\max}$ (i.e. for decreasing values).
(2) The bands of chaos start after the collision of an initial period-2 orbit with one of the endpoints of the dynamical core (θ_l, θ_r). One speaks in this case of a boundary collision bifurcation.
(3) The bifurcation points with respect to $x_{\min}$ occur before for $p_{\max} = 1$ than for $p_{\max} = 0.5$. The opposite happens with respect to $x_{\max}$. In either case we conclude that the setting $p_{\max} = 0.5$ is more robust than $p_{\max} = 1$.
(4) Last but not least, the bifurcation point with respect to $x_{\min}$ (respectively, $x_{\max}$) for $\alpha = 0.6$, $\beta = 0.4$ is greater (respectively, smaller) than for $\alpha = \beta = 1$. So the model with $\alpha = 0.6$, $\beta = 0.4$ is more stable with respect to both $x_{\min}$ and $x_{\max}$ than the original model (2.18).

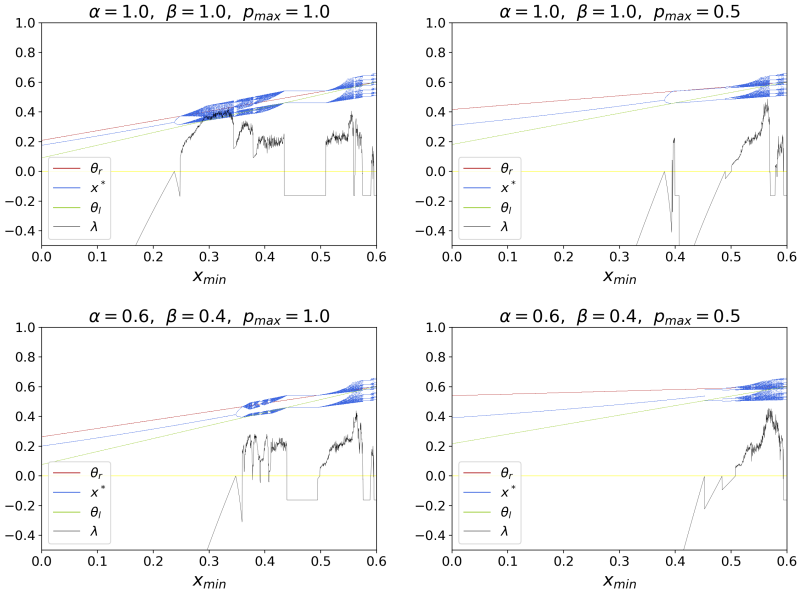

Fig. 3. Bifurcation diagram and Lyapunov exponents of the normalized average queue length x with respect to the parameter $x_{\min}$ for the values of α, β and $p_{\max}$ displayed at the top of the panels. The parameter $x_{\min}$ ranges in the interval $[0.2, 0.6)$. System parameters are given in (6.1). Other control parameters: $x_{\max} = 0.6$ and $w = 0.15$.

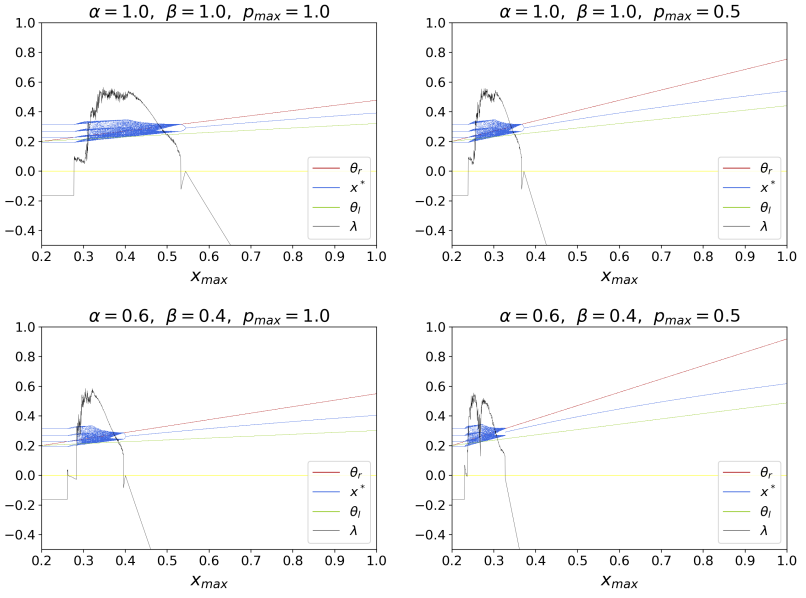

Fig. 4. Bifurcation diagram and Lyapunov exponents of the normalized average queue length x with respect to the parameter $x_{\max}$ for the values of α, β and $p_{\max}$ displayed at the top of the pannels. The parameter $x_{\max}$ ranges in the interval $(0.2, 1]$. System parameters are given in (6.1). Other control parameters: $x_{\min} = 0.2$ and $w = 0.15$.

Similar conclusions regarding the stability of x^* were obtained in [3] with respect to A_1, A_2 and w.

6.2. *Robustness domains*

Bifurcation diagrams are instrumental to find ranges of the control parameters that guarantee a stable dynamics. But this is not sufficient for our purposes. In this regard, an important feature in the design of AQM mechanisms is that small changes in the selected control parameters should not affect the stability of the system. In other words, the selected values of the control parameters (especially, α and β) should be also robust. As a result, robustness in the sense meant here, enables a stable operation under both changing system parameters and perturbations due to physical and numerical noise.

If bifurcations diagrams allow to find the stability ranges of different parameters, robustness domains allow to fine-tune the selection of control parameters for stability control under changing system parameters. In Fig. 5, we have scanned the interval $[0.002, 1.5] \times [0.002, 1.5]$ of the (α, β)-plane with precision $\Delta\alpha = \Delta\beta = 3.745 \times 10^{-3}$ (corresponding to a grid of 400×400 points). For each point (α, β) and $N = 1350$ (other system parameters as in (6.1)) we have calculated the bifurcation point of the averaging weight w_{bif}, see the left panel, for comparison with the reference value $N = 1850$ calculated in [3], see the right panel.

Regarding the robustness domains shown in the left panel of Fig. 5, we see that the second greatest domain corresponds to $0.15 \leq w_{\text{bif}} \leq 0.30$, and the third greatest to $0.30 \leq w_{\text{bif}} \leq 0.45$. Although the latter is smaller, settings of α, β in that domain (with a typical selection $\alpha = 0.5$, $\beta = 0.1$) leave an ample range $0 < w < 0.30$ for stable operation. Points (α, β) close to the boundaries of the robustness (same-color) domains should be avoided. Furthermore, comparison of the left panel with the right one shows that, as N increases, corresponding domains overlap, so one can find suitable selections of α and β for N ranging from 1350 to the reference number 1850. The observation that the robustness domains with

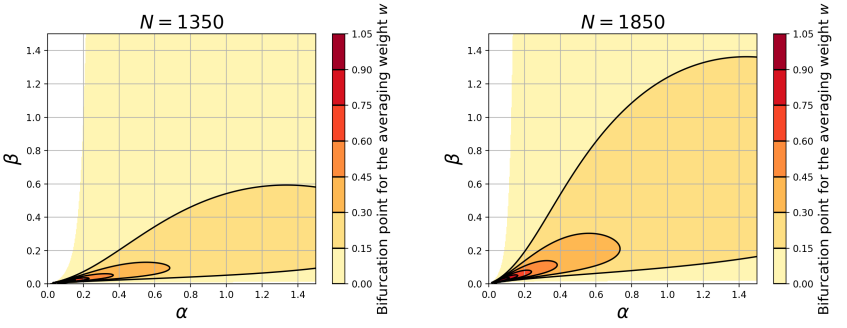

Fig. 5. (α, β)-parametric sweeps for the bifurcation point of the averaging weight w. Left panel: $N = 1350$; Right panel: $N = 1850$ (Source: [3]). Remaining system parameters are given in (6.1). Other control parameters: $x_{\min} = 0.2$, $x_{\max} = 0.6$.

$w_{\rm bif} \geq 0.15$ shrinks as N decreases agrees with the fact that a small number of connections (users) tends to disrupt the dynamics [19].

Of course, a thorough analysis is necessary for the actual design of an AQM algorithm. Our point here is that the decision on how to fix the control parameters is a compromise between the size of the robustness domains and stability ranges. We believe that the robustness domains are a handy tool for that endeavor.

7. Conclusion and Outlook

Deterministic models (mostly involving continuous time and differential equations) are typical of the traditional sciences, in which the evolution laws can be derived from first principles. Stochastic models are usually put in place in the opposite case, i.e. when the laws governing a process are unknown, hence a data-driven approach is the only option left. There are intermediate situations, though, where a deterministic description is possible but too complex to be of any practical use. A paradigmatic example is statistical mechanics, where the constituents move according to Newton's laws and, yet, a probabilistic approach (the observables being sample averages in phase space) has proved to be the best way to bridge the gap between the microscopic and the macroscopic scales. Other well-known example is symbolic dynamics, which is obtained by coarse-graining the state space of a discrete-time dynamical system.

In this paper, we proceeded the other way around: the outset was RED, a stochastic model to manage the incoming packet queue at a router's buffer. In this specific case, the uncertainty about the arrival times of the packets makes unfeasible a deterministic model. But if one contents oneself with aggregate information, which behaves in a more regular way, then the full power of a nonlinear model is available for prediction and control, showing that stochastic and deterministic models can complement each other advantageously. How to transit from the original RED model to a discrete-time dynamical formulation, where the states are average queue sizes at the buffer, was shown in Sec. 2. Since the resulting nonlinear version of RED turns out to be rather sensitive to the setting of control parameters (in particular, to the averaging weight w), the model was generalized in Sec. 3.1, Eq. (3.8), by replacing the original RED probability law (2.1) for packet dropping by a new one, Eq. (3.4), involving the beta distribution (3.5). This way, two new control parameters α and β are added; for $\alpha = \beta = 1$ one recovers the initial dynamical model (2.18). In Sec. 4, we gathered the main theoretical results on the global stability of x^*, the unique fixed point of the generalized RED model (3.8) if $A_1 < A_2 + x_{\max}$ (Theorem 4.1). Sufficient conditions were given in Theorem 4.6 for a monotone $f_{(\theta_l,\theta_r)}$, and in Theorems 5.1–5.2 for $f_{(\theta_l,\theta_r)}$ unimodal. Particular results for $f_{(\theta_l,\theta_r)}$ a $\cup$-convex function were derived in Sec. 5. The chaotic behavior of the RED dynamics, proved in Sec. 3.2, was quantified by bifurcation diagrams and Lyapunov exponents in Sec. 6.1, while the robustness of the fixed point under variations of the control parameters was discussed in Sec. 6.2.

The ultimate goal of the results reviewed and reported in this paper is the design of a congestion control algorithm for the data traffic on the Internet in a realistic environment, i.e. when some parameters of the communication network (notably, the number of users M and the round-trip time d) may vary over time. This was the main reason for studying convex mappings in Sec. 5 and parametric robustness in Sec. 6.2. The former provide an interesting trade-off between the implementation of the analytical constraints and practical controllability; the latter is crucial if the control parameters (especially α, β and w) have to be adjusted in real time. Congestion control under real conditions is the subject of work in progress.

Acknowledgments

We thank our reviewers for their constructive criticism. This material is partially based upon work supported by the Swedish Research Council under grant no. 2016-06596 while J. M. A. was in residence at Institut Mittag–Leffler in Djursholm, Sweden during the summer semester 2019. This work was also supported by Ministerio de Ciencia e Innovación, PID2019-108654GB-I00, Spain.

References

1. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, 1972).
2. R. Adams, Active queue management: A survey, *IEEE Commun. Surv. Tutorials* **15** (2013) 1425–1476.
3. J. M. Amigó, G. Duran, A. Giménez, O. Martí nez-Bonastre and J. Valero, Generalized TCP-RED dynamical model for Internet congestion control, *Commun. Nonlinear Sci. Num. Simul.* **82** (2020) 105075.
4. M. Arpaci and J. A. Copeland, An adaptive queue management method for congestion avoidance in TCP/IP networks, *IEEE Global Telecommun. Conf. (Globecom'00), Conf. Record (Cat. No. 00CH37137)*, **1** (2000) 309–315.
5. S. Athuraliya, V. H. Li, S. H. Low and Q. Yin, REM: Active queue management, *Teletraffic Science and Engineering* (Elsevier, 2001).
6. C. Brandauer, G. Iannaccone, C. Diot, T. Ziegler, S. Fdida and M. May, Comparison of tail drop and active queue management performance for bulk-data and Web-like Internet traffic, *Proc. Sixth IEEE Symp. Computers and Communications*, 2001, pp. 122–129.
7. K. Chin-Fu, S.-J. Chen, J.-M. Ho and R.-I. Chang, Improving end-to-end performance by active queue management, *19th Int. Conf. Advanced Information Networking and Applications (AINA'05), Vol. 1 (AINA papers)*, **2** (2005) 337–340.
8. M. Claypool, R. Kinicki and M. Hartling, Active queue management for Web traffic, *IEEE Int. Conf. Performance, Computing, and Communications*, 2004, pp. 531–538.
9. G. Duran, J. Valero, J. M. Amigó, A. Giménez and O. Martínez-Bonastre, Stabilizing chaotic behavior of RED, *2018 IEEE 26th Int. Network Protocols* (*ICNP*), Cambridge, UK, 2018, pp. 241–242.
10. G. Duran, J. Valero, J. M. Amigó, A. Giménez and O. Martinez-Bonastre, Bifurcation analysis for the Internet congestion, *2019 IEEE Conf. Computer Communications Workshops (INFOCOM)*, Paris, 2019, pp. 1073–1074.

11. H. A. El-Morshedy and E. Liz, Globally attracting fixed points in higher order discrete population models, *J. Math. Biol.* **53** (2006) 365–384.
12. S. Floyd and V. Jacobson, Random early detection gateways for congestion avoidance, *IEEE/ACM Trans. Network.* **1** (1993) 397–413.
13. S. Floyd, R. Gummadi and S. Shenker, Adaptive RED: An algorithm for increasing the robustness of RED's active queue management, https://www.icir.org/floyd/papers/adaptiveRed.pdf (2001), pp. 1–12.
14. J. P. Hespanha, S. Bohacek, K. Obraczka and J. Lee, Hybrid modeling of TCP congestion control. In M. D. Di Benedetto and A. Sangiovanni-Vincentelli (eds.), Hybrid Systems — Computation and Control: 4th International Workshop, HSCC 2001, Rome, Italy, March 2001, Proceedings (Lecture Notes in Computer Science, Vol. 2034), pp 291–304, Springer Verlag. ISBN 13: 9783540418665.
15. T. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Mon.* **82** (1975) 985–992.
16. M. Mathis, J. Semke, J. Mahdavi and T. Ott, The macroscopic behavior of the TCP congestion avoidance algorithm, *SIGCOMM Comput. Commun. Rev.* **27** (1997) 67–82.
17. J. Padhye, V. Firoiu, D. Towsley and J. Kurose, Modeling TCP Reno performance: a simple model and its empirical validation, *IEEE/ACM Trans. Network.* **8** (2000) 133–145.
18. P. Ranjan, E. H. Abed and R. J. La, Nonlinear instabilities in TCP-RED, *Proc. IEEE INFOCOM,* 2002, pp. 249–258.
19. P. Ranjan, E. H. Abed and R. J. La, Nonlinear instabilities in TCP-RED, *IEEE/ACM Trans. Network.* **12** (2004) 1079–1092.
20. A. N. Sharkovskiĭ, Coexistence of cycles of a continuous map of the line into itself, *Ukranian Math. J.* **16** (1964) 61–71.
21. W. Wang, H. Wang, W. Tang and F. Guo, A new approach to estimate RED parameters using function regression, *Int. J. Future Gen. Commun. Network.* **7** (2014) 103–118.