

# Mean-field analysis of a scaling MAC radio protocol

Illés Horváth<sup>1</sup> and Kristóf Attila Horváth<sup>2</sup> and Péter Kovács<sup>1</sup> and Miklós Telek<sup>1,2</sup>

<sup>1</sup> MTA-BME Information Systems Research Group

<sup>2</sup> Budapest University of Technology and Economics  
Department of Networked Systems and Services

**Abstract.** We examine the transient behavior of a positioning system with a large number of tags trying to connect to the infrastructure with an exponential backoff policy in case of unsuccessful connection. Using a classic mean-field approach, we derive a system of differential equations whose solution approximates the original process. Analysis of the solution shows that both the solution and the original system exhibits an unusual log-periodic behavior in the mean-field limit, along with other interesting patterns of behavior. We also perform numerical optimization for the backoff policy.

Keywords: population model, density dependent behavior, asymptotic limit, log-periodicity, positioning, ultra wideband, MAC, random access channel, exponential backoff

## 1 Introduction

### 1.1 Background and motivation

We examine the initial connecting process of an ultra wideband-based indoor positioning system with a large number of tags. During this initial connecting process, a random access channel is used with a classic exponential backoff policy for collision resolution. The considered performance problem is the initial connection time of the large number of tags at system (re)start, which might occur frequently in practice. One main goal of the paper is the analysis of the initial connection process by providing a mean-field approximation of its behaviour; an important advantage of the mean-field model over simulations is that the computational cost does not increase with the number of tags. The mean-field approximation allows us to perform numerical optimization for the backoff parameter, and we also introduce and numerically optimize an additional idea for improvement as well.

The rest of the paper is structured as follows. Section 1.2 contains the technical background of the system. In Section 1.3, we give a brief overview of the literature of random access channels, and in Section 1.4, a background on scaling limits.

Section 2 gives a proper mathematical description of the model. In Section 3.1, we derive the mean-field equations. In Section 3.2, the analysis of the process (both the original and the mean-field) is addressed, with several interesting observations and heuristics to help the understanding of the behavior exhibited by the system. Section 4 focuses on the mean time to connect, with numerical optimization of the parameters. In Section 5, we introduce and analyze an additional idea for improvement, and Section 6 concludes the work.

## 1.2 Technical background

Indoor positioning is one of the fastest growing areas today since technologies capable of positioning become more and more inexpensive [14]; meanwhile, the demand for location-based services is growing increasingly as well as the need for localization in the “Industry 4.0” conception. A lot of technologies can be used for positioning; however, most of those technologies, which are based on radio communication, can only use Received Signal Strength Indicator (RSSI) to determine distances, which may cause very inaccurate positioning. On the other hand, laser-based solutions are very accurate but too expensive. Ultrasound devices are also accurate, but their relatively small range is disadvantageous, hence limiting their applicability. An excellent candidate for indoor positioning is the Ultra Wideband (UWB) technology because it is relatively cheap, its range increment is in the order of tens of meters and its accuracy is high, because of time-of-flight measurements instead of RSSI detection.

UWB uses a very low power spectral density, but a large portion of the radio spectrum – by definition of Federal Communications Commission (FCC), the bandwidth exceeds 500 MHz or 20% of the center frequency [4]. Large bandwidth allows the emission of short pulses, therefore arrival times of packets can be determined accurately. Distances between devices are measured by a message exchanging method called two-way ranging (or by a more precise method called symmetric double-sided two-way ranging), which are defined in the IEEE 802.15.4 standard [16]. In an UWB-based positioning system, there are tags (to determine positions of moving components) and anchors (in well-known positions). After measuring distances from anchors, the position of a tag can be determined by trilateration. Depending on the application, the number of tags may vary from order 10 through order 100 [13] to order 10000 [17].

For accurate distance measurement, a long preamble is needed to start the packets in ranging method, which can become a long process. If the density of tags is high, a coordinated channel access is required, otherwise collisions could make it impossible to range. The coordinated channel is only accessible to the tags already connected to the system; a separate random access channel is used to connect the tags to the system.

A tree algorithm is often an excellent way to manage the collision resolution [3], for example, the ISO 18000-6 standard of RFID (Radio Frequency Identification) devices defines Type B mode, which uses a binary-tree collision-arbitration algorithm [6]. In the case of UWB communication, the problem with the tree algorithm is that empty slots must be distinguishable from timeslots in which

collisions occur. This is not possible, since the power spectral density of UWB communication is below the noise level of conventional radio services, thus carrier detection is impossible.

### 1.3 Random access channels

A classic approach to resolve collisions in random access channels is by using a backoff algorithm - in each timeslot, each tag tries to connect with a given probability. In case of a collision, the colliding tags decrease this probability for later timeslots. A widely-used approach to decreasing the probabilities is the use of a “contention window”: each tag selects a window length and a uniform random timeslot within the window for the next trial [10, 11]. In this approach, the backoff corresponds to increasing the window size. A widely-used scheme is binary exponential backoff (BEB), where the size of the window is doubled in each step. Increasing the window size exponentially corresponds to decreasing the connection probabilities by the same factor (see Figure 1 and formula (3) on page 3 of [11]).

Exponential backoff algorithms have been usually examined in the stationary setting, where tags continuously arrive and connect. For the stability condition for such systems, see [10], and for the performance analysis, see [11]. Since then, several minor improvements in the algorithm have also been introduced, see e.g. [15].

In some versions of the algorithm, the number of retransmissions is maximized; after reaching the maximum, the tag is dropped (see *EB-M* in [11]). While this considerably simplifies the analysis and improves the performance of the system, this is only applicable if the number of tags simultaneously present in the system is low. This is a valid assumption for stable stationary systems, but might not hold for transient scenarios.

Our main concern is the transient analysis when the system starts with a large number of tags. This may be the case for example when the system is rebooted after a failure, and many tags try to connect initially. The performance of the system in the transient case is very different from the stationary case, and, to the best of our knowledge, have not been examined in depth in the literature. In this scenario, without prior knowledge on the number of tags trying to connect, setting a maximum on the number of retransmissions is not feasible. However, a similar but viable idea instead is introduced and analyzed in Section 5.

### 1.4 Scaling limits

Population models have been examined widely, with applications ranging from biology through queuing to physics (among other topics) [12]. To understand the behavior of large systems, one usually needs to understand the correct scaling for population, time or space (whichever is applicable to a specific model), along with the scaling limit of the process.

Mean field theory has been applied to identify the scaling limit of various queuing models. The classic result of Kurtz [9] identifies the scaling limit of the

evolution of a density-dependent Markov population process as the solution of a deterministic system of ordinary differential equations (ODEs). Kurtz involves population scaling, but no time scaling (apart from the natural fact that the total number of events over a given time period is proportional to the population size). Kurtz and many other similar results are examples of the classic mean field approach: by approximating the behavior of each individual in the population by its expectation, the system converges to a deterministic mean field limit.

Many physical systems, most notably models of random walks in random environment and interacting particle systems, involve the scaling of time and space as well. Examples include the simple exclusion process [7], where the trajectory of a tagged particle is tracked in a space including other particles. The scaling limit of such systems is often random (e.g. Brownian motion), so the mean field approach does not apply.

Since then, many other results and applications have been developed; see [5] for the recent theory on the convergence of Markov processes and [1] for the mean-field theory for scale-free random networks. For a recent book that deals with random walks in random environment and interacting particle systems, see [8].

One interesting phenomenon that occurs in very few systems is log-periodicity. Usually, convergence to the scaling limit holds as the system size increases. For log-periodic systems, this is not the case: as the size increases, instead of convergence, the behavior changes periodically with the logarithm of the system size. In [2], the authors examine the hitting times of levels for a certain type of random walk on infinite random trees. In that setup, log-periodicity occurs due to long periods of time spent in so-called “traps” which remain visible in the mean-field limit. In [18], log-periodic behavior in time was found for random walks on randomly diluted cubic lattices.

Interestingly, the simple protocol described in Section 2 exhibits both scaling and log-periodicity.

## 2 Model description

### 2.1 Protocol

As discussed in Section 1, the analyzed UWB system contains moving tags and anchors in fixed positions. For simplicity, we assume that every device is in the range of each other. The positioning and the normal communication between the tags and the anchors are controlled; however, this controlled structure contains random access slots, when unconnected tags can connect to the system.

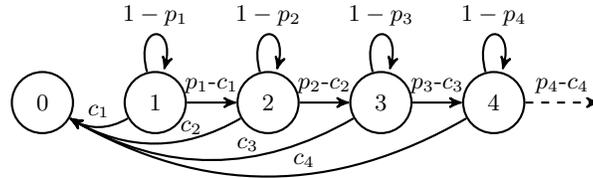
If a tag is in the unconnected state, it is waiting for the next connection slot and then transmits a packet with probability  $p$ . (When the tag goes to the unconnected state, probability  $p$  is initialized to 1.) If the packet is not acknowledged, then the tag reduces  $p$  according to a backoff strategy, e.g., it halves the probability.

If the connection packet is received and hence acknowledged, then the tag becomes connected, therefore it leaves the connection process and gets coordinated channel access eventually.

The performance of the detailed connection protocol is critical when the system is rebooting and thus every tag switches to the unconnected state. In this case a large number of devices try to access the channel simultaneously.

## 2.2 Mathematical model

The protocol described in Section 2.1 translates to a probabilistic model, which we examine in detail here. Initially, there are  $N$  competing components (referred to as tags in Section 1.2) in a discrete time setting. In each step (time slot), each competing component is trying to connect with some probability  $p_i$  where  $i$  corresponds to the number of times backoff has occurred for the given component. In other words, the competing components can be arranged into classes indexed by  $i$ , where each component in class  $i$  is trying to connect with probability  $p_i$  in each step. For simplicity, we omit the trivial class with  $p_i = 1$ , and assume that the initial class is  $i = 1$  with  $p_1 < 1$ . If a component successfully connects (that is, no other components tried to connect), it moves to class 0, which means that it is connected and it does not compete for channel access any more. If a component tries to connect but fails, it moves to class  $i + 1$  instead. The behavior of the system is displayed in Figure 1.



**Fig. 1.** State transitions of a single user;  $p_i$  are constant,  $c_i$  depend on other users

If  $N_i(n)$  denotes the number of components in class  $i$  at time  $n$ , then  $\{N_i(n) : i \geq 0\}$  is a discrete time Markov chain. One step of the evolution of this Markov chain can be generated in the following way:

- for each class  $i$ , the number of components trying to connect is  $C_{i,n} \sim \text{Binomial}(N_i(n), p_i)$  (independently for each class);
- if  $\sum_{j=1, j \neq i}^{\infty} C_{j,n} = 0$  and  $C_{i,n} = 1$ , then a single component from class  $i$  connects, so

$$N_i(n+1) = N_i(n) - 1, \quad N_j(n+1) = N_j(n) \quad (j \neq i); \quad (1)$$

- if  $\sum_{j=1}^{\infty} C_{j,n} \neq 1$ , then no component connects, and all components that tried to connect transition to the next class, so

$$N_i(n+1) = N_i(n) - C_{i,n} + C_{i-1,n} \quad \forall i. \quad (2)$$

The probability that a given component from class  $i$  connects can be calculated as

$$c_i(n) = p_i(1 - p_i)^{N_i(n)-1} \prod_{j=1, j \neq i}^{\infty} (1 - p_j)^{N_j(n)} = \frac{p_i}{1 - p_i} \prod_{j=1}^{\infty} (1 - p_j)^{N_j(n)}, \quad (3)$$

since

$$\begin{aligned} P(\text{the component tries to connect in class } i) &= p_i, \\ P(\text{no other component tries to connect in class } i) &= \\ (1 - p_i)^{N_i(n)-1}, \\ P(\text{no component tries to connect in any other class}) &= \\ \prod_{j=1, j \neq i}^{\infty} (1 - p_j)^{N_j(n)}. \end{aligned}$$

Note that the different classes are still coupled via the condition  $\sum_{i=1}^{\infty} C_{i,n} = 1$ , which is necessary for a successful connection.

The main question of interest is the behavior of the system for large values of  $N$ . How long does it take for all components to connect? Does mean-field convergence hold for the ratio of connected components  $N_0(n)/N$ , or, equivalently, the ratio of competing components in the system  $1 - N_0(n)/N$ ; that is, is there a deterministic function  $z(t)$  going from 1 to 0 such that, with a proper scaling of time  $n(t)$ ,

$$\sum_{i=1}^{\infty} \frac{N_i(n(t))}{N} = 1 - \frac{N_0(n(t))}{N} \rightarrow z(t) \quad \text{as } N \rightarrow \infty \quad (4)$$

holds? What can we say about the classes  $N_i$ ,  $i \geq 1$ ? We are also interested in practical questions: which backoff policy performs best according to some performance measures (e.g. average time to connect)?

In the present paper, we focus on the exponential backoff policy, when  $p_i = \gamma^{-i}$  for some  $\gamma > 1$ ; this model offers a surprisingly rich and unexpected behavior.

### 3 Mean-field approximation

#### 3.1 Deriving the mean-field equations

From the formulas (1)–(2), the expected change in  $N_i$  can be calculated as

$$\begin{aligned} E(N_i(n+1) | \{N_j(n), j \geq 0\}) &= \\ N_i(n) - p_i N_i(n) + (p_{i-1} - c_{i-1}(n)) N_{i-1}(n); \end{aligned} \quad (5)$$

on the right hand side, the term  $-p_i N_i(n)$  corresponds to the average number of components in class  $i$  trying to connect (these leave class  $i$  regardless of

whether the connection is successful or not), the term  $(p_{i-1} - c_{i-1}(n))N_{i-1}(n)$  corresponds to the average number of components arriving in class  $i$  from class  $i-1$ :  $p_{i-1}N_{i-1}(n)$  is the average number of components leaving class  $i-1$ , and  $c_{i-1}(n)N_{i-1}(n)$  is the average number of components from among them who managed to connect (and thus end up in class 0 instead of class  $i$ ).

Technically, (5) is valid only for  $i > 1$ , while for  $i = 1$ ,

$$E(N_1(n+1)|\{N_j(n), j \geq 0\}) = N_1(n) - p_1 N_1(n) \quad (6)$$

holds instead. For easier notation, we use only the form (5) throughout the rest of the paper, understood to include (6) as well.

Using a classic mean-field approach, we define the process  $x_i(n) = \frac{N_i(n)}{N}$  (which is normalized by  $N$  compared to  $N_i(n)$ ) with evolution defined according to the expected behavior as in (5):

$$x_i(n+1) = x_i(n) - p_i x_i(n) + (p_{i-1} - c_{i-1}(n))x_{i-1}(n). \quad (7)$$

(7) is general in the sense that so far, we have not used the assumption that the backoff policy is exponential. From this point on in the calculations, we always assume that  $p_i = \gamma^{-i}$  (with  $\gamma > 1$ ).

In order to derive the mean-field limit of  $x_i(n)$ , we denote the integer and the fraction part of  $\log_\gamma N$  by  $L = \lfloor \log_\gamma N \rfloor$  and  $\alpha = \{\log_\gamma N\}$ , respectively, and consequently  $N = \gamma^{\alpha+L}$ .

For  $i > 1$ ,

$$x_i(n+1) = x_i(n) - \gamma^{-i} x_i(n) + (\gamma^{-(i-1)} - c_{i-1}(n))x_{i-1}(n), \quad (8)$$

where

$$\begin{aligned} c_i(n) &= \prod_{j=1}^{\infty} (1 - \gamma^{-j})^{N x_j(n)} \frac{\gamma^{-i}}{1 - \gamma^{-i}} \\ &= \exp \left( \sum_{j=1}^{\infty} N x_j(n) \log(1 - \gamma^{-j}) \right) \frac{\gamma^{-i}}{1 - \gamma^{-i}} \\ &\simeq \exp \left( -N \sum_{j=1}^{\infty} x_j(n) \gamma^{-j} \right) \frac{\gamma^{-i}}{1 - \gamma^{-i}} \simeq \exp \left( -N \sum_{j=1}^{\infty} x_j(n) \gamma^{-j} \right) \gamma^{-i}, \end{aligned} \quad (9)$$

where we used  $\log(1 - \gamma^{-j}) \simeq -\gamma^{-j}$  and  $(1 - \gamma^{-j}) \simeq 1$ . Note that these approximations are valid only if  $j$  is large, but according to Remark 3 later, those are the states where the process shows interesting behavior, while the terms corresponding to states where  $j$  is small, vanish. Then

$$\begin{aligned} x_i(n+1) - x_i(n) &\simeq \\ &- \gamma^{-i} x_i(n) + \left( 1 - \exp \left( - \sum_{j=1}^{\infty} \gamma^{\alpha+L} \gamma^{-j} x_j(n) \right) \right) \gamma^{-(i-1)} x_{i-1}(n). \end{aligned}$$

Up to this point we rescaled the ‘size’ of the process  $x_i(n) = \frac{N_i(n)}{N}$ , next we rescale its ‘speed’ by introducing  $v_i(t) = x_i(n)$ , where  $n = tN$ . For  $v_i(t)$ , we write

$$v_i(t + \gamma^{-(\alpha+L)}) - v_i(t) \simeq \gamma^{-(\alpha+L)} \frac{dv_i(t)}{dt} \simeq -\gamma^{-i} v_i(t) + \left( 1 - \exp \left( - \sum_{j=1}^{\infty} \gamma^{\alpha+L} \gamma^{-j} v_j(t) \right) \right) \gamma^{-(i-1)} v_{i-1}(t).$$

In order to ease the technical description we also shift the index by  $L$ , that is,  $w_{i-L}(t) = v_i(t)$ , and write

$$\gamma^{-(\alpha+L)} \frac{dw_{i-L}(t)}{dt} \simeq -\gamma^{-i} w_{i-L}(t) + \left( 1 - \exp \left( - \sum_{j=1}^{\infty} \gamma^{\alpha+L} \gamma^{-j} w_{j-L}(t) \right) \right) \gamma^{-(i-1)} w_{i-L-1}(t).$$

This last step is motivated by the fact that the interesting behavior (the dominant part of transitions to class 0) occurs around class  $L$  and by this index shifting it gets to be around index zero independent of the size of the population.

Rearranging the  $\gamma^L$  factor in the last expression gives

$$\frac{dw_{i-L}(t)}{dt} \simeq -\gamma^\alpha \gamma^{-(i-L)} w_{i-L}(t) + \left( 1 - \exp \left( - \sum_{j=1}^{\infty} \gamma^\alpha \gamma^{-(j-L)} w_{j-L}(t) \right) \right) \gamma^\alpha \gamma^{-(i-L-1)} w_{i-L-1}(t),$$

or, after re-indexing, simply

$$\frac{dw_i(t)}{dt} = -\gamma^\alpha \gamma^{-i} w_i(t) + \left( 1 - \exp \left( - \sum_{j=-L+1}^{\infty} \gamma^\alpha \gamma^{-j} w_j(t) \right) \right) \gamma^\alpha \gamma^{-(i-1)} w_{i-1}(t). \quad (10)$$

where we write an equality in order to get a proper differential equation. The initial condition corresponds to  $x_1(0) = 1$ ,  $x_i(0) = 0$ ,  $i > 1$ , which translates to  $w_{1-L}(0) = 1$ ,  $w_i(0) = 0$ ,  $i > 1 - L$ .

*Remark 1.* The equation (10) can be written in an equivalent form by denoting  $y_i(t) = \gamma^\alpha w_i(\gamma^{-\alpha} t)$  to get

$$\frac{dy_i(t)}{dt} = -\gamma^{-i} y_i(t) + \left( 1 - \exp \left( - \sum_{j=-L+1}^{\infty} \gamma^{-j} y_j(t) \right) \right) \gamma^{-(i-1)} y_{i-1}(t),$$

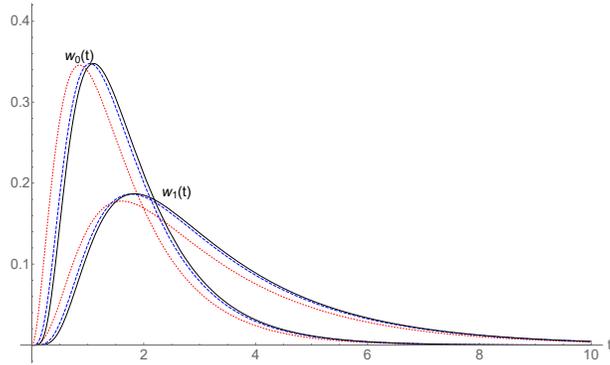
with the initial condition  $y_{-L+1}(0) = \gamma^\alpha, y_i(0) = 0, i > -L + 1$ . In this case,  $\gamma^\alpha$  is present in the initial condition instead of the differential equation.

That said, we continue with the analysis of equation (10). The parameters  $\alpha$  and  $L$ , which are related to the size of the population, are still included in the equations, but now in an intrinsically different manner. We can evaluate the limit as  $L$  tends to infinity such that  $\alpha$  remains present in the model description. Experimental analysis shows that as  $L \rightarrow \infty$  through integer values (and  $\alpha$  and  $\gamma$  are fixed), the solution of (10) converges to a limit

$$\lim_{L \rightarrow \infty} w_i(t) = z_i(t) = z_i(\gamma, \alpha, t), \quad -\infty < i < \infty. \quad (11)$$

( $\alpha$  and  $\gamma$  will often be omitted from the notation).

Figure 2 displays the convergence of the functions  $w_i(t)$  as  $L$  increases;  $w_0(t)$  and  $w_1(t)$  for  $L = 2$  correspond to the dotted red lines, while  $w_0(t)$  and  $w_1(t)$  for  $L = 4$  correspond to the dashed blue lines. The continuous black lines are the limit functions  $z_0(t)$  and  $z_1(t)$  (other parameters are  $\alpha = 0$  and  $\gamma = 2$ ). The convergence is rather fast, with  $w_i(t)$  close to  $z_i(t)$  already for  $L = 4$  (which corresponds to a population size of  $N = 2^4 = 16$ ).



**Fig. 2.** Convergence of  $w_0(t)$  and  $w_1(t)$  when  $\alpha$  is fixed and  $L \rightarrow \infty$

One may heuristically think of  $z_i(t)$  as the solution of

$$\frac{dz_i(t)}{dt} = -\gamma^\alpha \gamma^{-i} z_i(t) + \left( 1 - \exp \left( - \sum_{j=-\infty}^{\infty} \gamma^\alpha \gamma^{-j} z_j(t) \right) \right) \gamma^\alpha \gamma^{-(i-1)} z_{i-1}(t). \quad (12)$$

The main issue with (12) is that it has no meaningful initialization at  $t = 0$ . However, for any  $t_0 > 0$ , we may obtain an initialization at  $t_0$  as the limit of

$w_i(t_0)$  as  $L \rightarrow \infty$ , and then the process  $z_i(t)$  does evolve according to (12) for  $t > t_0$ . We will address calculating  $w_i(t_0)$  for small  $t_0$  later, see Remark 4.

A numerical solution of (12) (or (10)) may be obtained by solving the following finite system for  $\hat{z}_i(t)$ ,  $-m \leq i \leq M$ , where  $m$  and  $M$  are selected sufficiently large:

$$\begin{aligned} \frac{d\hat{z}_i(t)}{dt} = & -\gamma^\alpha \gamma^{-i} \hat{z}_i(t) + \\ & \left( 1 - \exp \left( - \sum_{j=-m}^M \gamma^\alpha \gamma^{-j} \hat{z}_j(t) \right) \right) \gamma^\alpha \gamma^{-(i-1)} \hat{z}_{i-1}(t), \quad (13) \\ \hat{z}_{-m}(0) = & 1, \quad \hat{z}_i(0) = 0, \quad i = -m + 1, \dots, M. \end{aligned}$$

Standard numerical solvers are feasible for (13). The roles of  $m$  and  $M$  are different:  $m$  effectively corresponds to  $L$  (actually, Figure 2 was obtained by setting  $m = L - 1$ ), so the error made by choosing a finite  $m$  corresponds to the difference between  $z_i(t)$  and  $w_i(t)$ . The choice of  $M$ , on the other hand, serves to make the system finite-dimensional by truncating high-index terms which are negligible in practice anyway (see also Remark 2 later).

Technically, all numerical solutions for  $z_i(t)$  throughout the paper are obtained from (13), but for simplicity, we will only use the notation  $z_i(t)$  (except when addressing the difference between  $z_i(t)$  and  $\hat{z}_i(t)$  explicitly).

### 3.2 The system behavior and its mean-field description

In this section we compare the simulations of the original stochastic model  $\{N_i(n)\}$  with its mean-field counterpart  $z_i(\gamma, \alpha, t)$ , that is, we check the validity of

$$z_i(\gamma, \alpha, t) \approx N_{L+i}(n)/N \quad \text{where } N = \gamma^{\alpha+L} \text{ and } n = Nt. \quad (14)$$

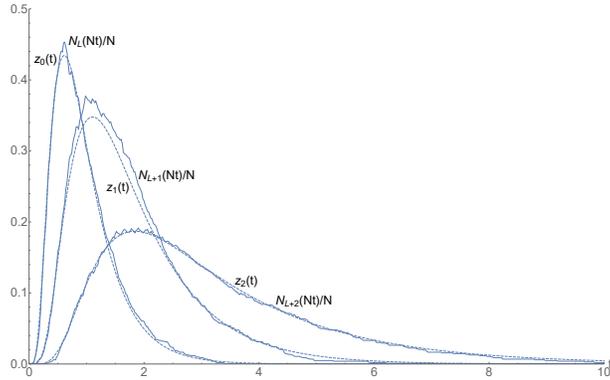
If (14) holds, then for large population sizes we can compute performance measures of interest based on  $z_i(\gamma, \alpha, t)$  (as it is done in the next section) instead of the solution of the Markov chain  $N_i(n)$  whose state space has prohibitive size.

Figure 3 displays simulation results against the numerical solution of  $z_i(t)$ :

- the smooth dashed lines are the numerical solution for  $z_0(t)$ ,  $z_1(t)$  and  $z_2(t)$  with parameters  $\gamma = 2$  and  $\alpha = 0$ ;
- the jagged lines are simulation results for  $N_L(Nt)/N$ ,  $N_{L+1}(Nt)/N$  and  $N_{L+2}(Nt)/N$  for the original system for  $N = 2^{10}$  and  $\gamma = 2$ .

(14) means the following scaling for the original process:

- (align with expectations) the population size scales with  $N$ ,
- time scales with  $N$ , and
- the space (that is, the indices of classes where the bulk of the process is concentrated) is shifted by  $L = \lfloor \log_\gamma N \rfloor$ , but not scaled.



**Fig. 3.** Simulation for  $N_{L+i}(Nt)/N$  versus numerical solution for  $z_i(t)$  for  $i = 0, 1, 2$  (parameters are  $N = 2^{10}$ ,  $\gamma = 2$ ,  $L = 10$ ,  $\alpha = 0$ )

From this point on, the presentation of the analysis is often two-fold: we examine the behavior of  $z_i(t)$ , and for each property identified, we provide a corresponding heuristic explanation of the same behavior for the original stochastic model  $N_i(n)$ .

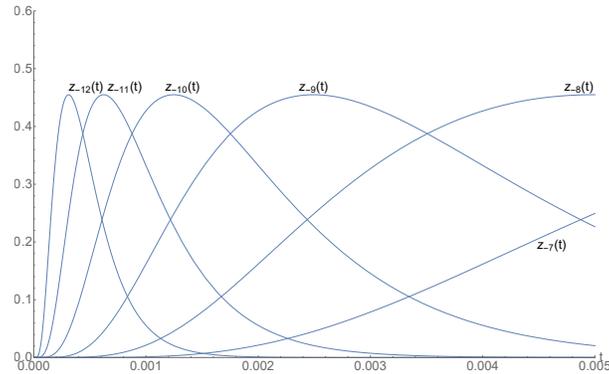
Numerical solution of (10) shows the following behavior for large  $N$  (i.e. large  $L$ ):

- (a) for values of  $i$  which are considerably smaller than 0,  $z_i(t)$  has a local maximum at a time of order  $\gamma^{-i}$ , after which the function decays rapidly (exponentially with very large rate);
- (b) for values of  $i$  around 0,  $z_i(t)$  has a local maximum at a time of order 1, after which it decays exponentially (at a fixed rate of order 1);
- (c) for values of  $i$  considerably larger than 0,  $z_i(t)$  remains very close to 0 all the time; that said, each  $z_i(t)$  has a rather slow exponential decay (with the decay rate going to 0 as  $i \rightarrow \infty$ ).

The corresponding behavior of the original stochastic model:

- (a) the bulk of the components “run through” the early classes very rapidly, spending order  $N/\gamma^i$  steps in class  $i$  ( $i = 1, 2, \dots$ ) in average;
- (b) the bulk of the components spend a number of steps proportional to  $N$  near the class  $L$ ;
- (c) the majority of the components never get to a level much higher than  $L$ .

We refer to item (a) of the above behavior as *early rapid transition*. It is demonstrated in Figure 4, which depicts  $z_i(t)$  for values of  $i$  considerably smaller than 0. Note that  $z_i(t)$  has its local maximum near  $\gamma^{-i} = 2^{-i}$ . We also note that during the early rapid transition phase, essentially no components manage to connect.



**Fig. 4.** Early rapid transition:  $z_i(t)$  for values of  $i$  considerably smaller than 0 ( $\alpha = 0$  and  $\gamma = 2$ )

*Remark 2.* The width of the “event window” around 0 is fixed; that is, the interval of classes from which the majority of successful transitions take place, item (b), does not increase as  $N \rightarrow \infty$ . In other words, for the original model, this means that the bulk of the components are contained in the same number of classes regardless how large  $N$  is. This also means that the choice of  $M$  in (13) for the numerical solution is “absolute”: since the event window does not scale with  $N$ , any error introduced by using a finite  $M$  does not scale either, so the error can be set arbitrarily small by setting  $M$  large enough, regardless of  $N$ . Numerical investigations indicate  $M = 10$  to be sufficient.

*Remark 3.* During the approximation of  $c_i$  in Section 3.1, we used the approximation  $1 - \gamma^{-i} \simeq 1$ , which was valid only if  $i$  was large. For small values of  $i$ , this introduces a constant multiplicative error in the term corresponding to  $c_i$ . However, for small values of  $i$ ,  $w_i(t)$  (along with  $z_i(t)$ ) decays rapidly, so it only exhibits nontrivial behavior for very small values of  $t$ , and at those values of  $t$ , the exponential term in (9) is very small. This means that the actual error introduced by the approximation of  $c_i$  is negligible when  $N$  is large. For the original stochastic system, this corresponds to the fact that during the very early stages of the process, when the bulk of the components are in the early classes, the value of  $c_i$  is extremely small and so very few components manage to connect, rendering any error in the approximation of  $c_i$  overall irrelevant.

*Remark 4.* Neglecting the exponential term in (10) entirely, we get the equations

$$\begin{aligned} \frac{d\hat{w}_i(t)}{dt} &= -\gamma^\alpha \gamma^{-i} \hat{w}_i(t) + \gamma^\alpha \gamma^{-(i-1)} \hat{w}_{i-1}(t), \\ \hat{w}_{-L+1}(0) &= 1, \quad \hat{w}_i(0) = 0, \quad i > -L + 1. \end{aligned} \quad (15)$$

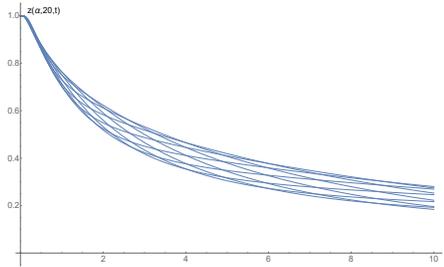
This is a simple linear system with an explicit analytical solution that approximates (10) nicely for small values of  $t$ . (Of course, for larger values of  $t$ , the behavior of  $w_i(t)$  and  $\hat{w}_i(t)$  diverge eventually.)

For different values of  $\alpha$ , the functions  $z_i(\gamma, \alpha, t)$  are different. This means that the original system  $N_i(t)$  exhibits a so-called *log-periodic* behavior: that is, as  $N \rightarrow \infty$ ,  $N_{L+i}(Nt)/N$  does not converge in general, but if  $N = \gamma^{\alpha+L}$  where  $L \rightarrow \infty$  and  $\alpha$  is fixed, then  $N_{i+L}(Nt)/N$  converges to  $z_i(\gamma, \alpha, t)$ .

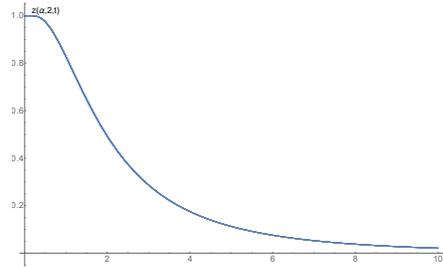
We further introduce

$$z(\gamma, \alpha, t) = \sum_{i \in \mathbb{Z}} z_i(\gamma, \alpha, t),$$

the ratio of components still in the system at time  $t$ ; this is the original limit we were looking for in (4) in Section 2. Figure 5 displays the functions  $z(20, \alpha, t)$  for  $\gamma = 20$  and  $\alpha = 0, 1/10, \dots, 9/10$  ( $z(\gamma, 0, t)$  is thick). Note that the “amplitude” of this periodic behavior is much smaller for smaller values of  $\gamma$ ; Figure 6 displays the functions  $z(\gamma, \alpha, t)$  for  $\gamma = 2$  and  $\alpha = 0, 1/10, \dots, 9/10$  which turn out to be much closer than for  $\gamma = 20$  (although still slightly different).



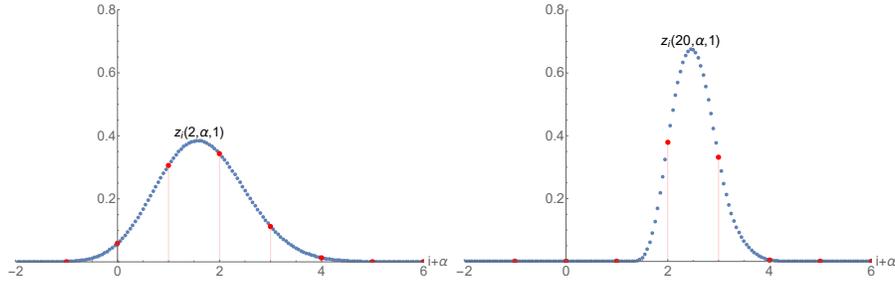
**Fig. 5.** The functions  $z(\gamma, \alpha, t)$  for  $\gamma = 20$  and  $\alpha = 0$  (thick line),  $1/10, \dots, 9/10$



**Fig. 6.** The functions  $z(\gamma, \alpha, t)$  for  $\gamma = 2$  and  $\alpha = 0, 1/10, \dots, 9/10$

Nevertheless, the individual functions  $z_i(t)$  still very much depend on  $\alpha$  even for small values of  $\gamma$ ; Figure 7 displays the values of the functions  $z_i(\gamma, \alpha, t)$  for  $\gamma = 2$  at a given point in time ( $t = 1$ ). The value of  $\alpha$  is incorporated into the figure by a shift by  $\alpha$  to the right; for example, the values  $z_i(\gamma, 0, 1), i \in \mathbb{Z}$  are positioned exactly above integers (and are marked with columns), while the values  $z_i(\gamma, 1/20, 1), i \in \mathbb{Z}$  are positioned above the points  $\{i + 1/20, i \in \mathbb{Z}\}$  and so on. Displaying the values this way makes some sort of continuous background function appear. A detailed examination shows that it is not symmetric and the shape of the function changes with  $t$  and converging to 0 as  $t \rightarrow \infty$ .

If we keep  $t = 1$  fixed and examine the background function displayed in Figure 7 for other values of  $\gamma$ , it turns out that for higher values of  $\gamma$  it is more concentrated, as seen in Figure 8 which displays the values  $z_i(\gamma, \alpha, t)$  for  $\gamma = 20$  at  $t = 1$  (similarly to Figure 7). This phenomenon has a heuristic explanation for the original stochastic model: when a component transitions to the next class, this transition means a more drastic change in the behavior of the component for higher values of  $\gamma$ ; as a result, for higher values of  $\gamma$ ,



**Fig. 7.** The values  $z_i(2, \alpha, 1)$  for  $\alpha = 0, 1/20, \dots, 19/20$  **Fig. 8.** The values  $z_i(20, \alpha, 1)$  for  $\alpha = 0, 1/20, \dots, 19/20$

the bulk of the components will be concentrated in fewer classes, making the background continuous function more concentrated as well. This also explains the phenomenon that the “amplitude” of the periodic behavior is larger for higher values of  $\gamma$ : for a background function as concentrated as the one in Figure 8, changing the value of  $\alpha$  creates a more drastic change in the distribution of the “mass” over the classes (for example, in Figure 8 for  $\alpha = 0$  (red columns), the bulk of the mass is contained in two classes  $i = 2$  and  $i = 3$ ; however, for  $\alpha = 1/2$ , the bulk of the mass is contained in a single class displayed over the point 2.5).

*Remark 5.* Denote

$$\lambda(t) = \sum_{i=-\infty}^{\infty} \gamma^{-i} z_i(t), \quad z(t) = \sum_{i=-\infty}^{\infty} z_i(t).$$

Based on (12), for  $z(t)$  and  $\lambda(t)$  we get

$$\frac{dz(t)}{dt} = -\lambda(t) + (1 - \exp(-\lambda(t))) \lambda(t) = -\lambda(t) \exp(-\lambda(t)). \quad (16)$$

$z(t)$  and  $\lambda(t)$  are related with the elements of the original model defined in Section 2 as

$$z(t) \simeq \sum_{i=1}^{\infty} N_i(Nt)/N, \quad \text{and} \quad \lambda(t) \simeq \sum_{i=1}^{\infty} p_i N_i(Nt), \quad (17)$$

where  $\sum_{i=1}^{\infty} p_i N_i(n)$  is the expected number of components trying to connect. We assume that the number of components trying to connect has a distribution very close to Poisson with parameter  $\sum_{i=1}^{\infty} p_i N_i(n)$ . It is supported by the following facts: if  $X_i \sim \text{Binom}(N_i(n), p_i)$  where  $N_i(n)$  is large and  $p_i$  is small, then  $X_i$  is very close to Poisson with parameter  $p_i N_i(n)$ , and the sum of independent Poisson variables is also Poisson with the parameters adding up. It follows that the number of components trying to connect has a random distribution very

close to Poisson with parameter  $\sum_{i=1}^{\infty} p_i N_i(n) \simeq \lambda(t)$ . To obtain (16) we just have to note that the total number of components in the system decreases iff there is exactly one component trying to connect, and indeed,

$$P(X = 1) = \lambda(t) \exp(-\lambda(t)) \quad \text{for } X \sim \text{Poisson}(\lambda(t)), \quad (18)$$

explaining (16).

*Remark 6.* For small values of  $t$ , the value of  $\lambda(t)$  is very large, but it decreases rapidly as the process progresses. During the early rapid transition, the value of  $\lambda(t)$  is very high and the function  $\lambda(t) \exp(-\lambda(t))$  is very close to 0, so practically no components connect. A visible portion of the components starts to connect when  $\lambda(t)$  reaches the region of 1, and the bulk of the components connect while  $\lambda(t)$  is in this region. Then, after the majority of the components have connected, the value of  $\lambda(t)$  gets closer and closer to 0. The maximum of this function is

$$\max_{\lambda} \{\lambda \exp(-\lambda)\} = e^{-1} \quad \text{and it is obtained at } \lambda = 1. \quad (19)$$

While the exact time  $t$  when the value of  $\lambda(t)$  is equal to 1 depends on  $\gamma$  and  $\alpha$ , the actual value at the local maximum is always  $e^{-1}$ . This implies that

$$\max_t |z'(t)| = e^{-1} \quad (20)$$

for any setup of the parameters. This behavior also holds for the original stochastic process in the sense that the largest average rate (throughput) with which the components connect throughout the process is  $e^{-1}$  connections/time step. Actually, (20) holds in other settings as well, e.g. for a stable, stationary system with arrivals as in [10], formula (27).

*Remark 7.* We analyze the tail of  $z(t)$ . When  $t$  is large,  $z(t)$  is relatively small, and most of the contribution in  $z(t) = \sum_{i=-\infty}^{\infty} z_i(t)$  comes from larger index terms. For the original model, this means that most of the components have managed to connect, and the remaining components are concentrated in classes with a larger index. See also part (c) of the description of the behavior provided early in Section 3.2. When  $t$  is large,  $\lambda(t)$  is small, meaning that there is essentially no interaction between the components anymore: as we are nearing the end of the connection process, only few components remain, each trying to connect with a very small probability, and the probability of collision is negligible. Without interaction, each  $z_i(t)$  just decays exponentially with the rate going to 0 as  $i \rightarrow \infty$ . Overall, this means that the decay of  $z(t) = \sum_{i=-\infty}^{\infty} z_i(t)$  is superexponential.

## 4 Analysis of the mean time to connect

In this section we study the mean (scaled) time it takes to connect for a random component. For the system (10), the mean of the (scaled) connection time can

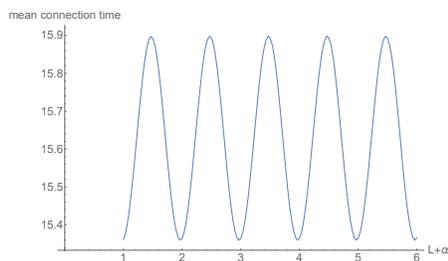
be obtained as

$$m_w(\gamma, \alpha, L) = \int_{t=0}^{\infty} \sum_{i=-L+1}^{\infty} w_i(t) dt, \quad (21)$$

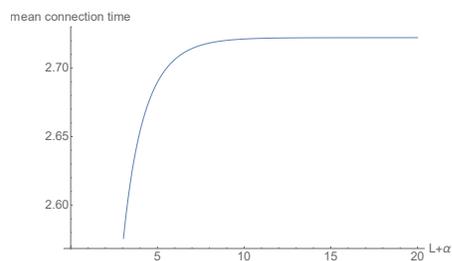
where  $w(t) = \sum_{i=-L+1}^{\infty} w_i(t)$  is the ratio of components still in the system. In accordance with (11), we also define

$$m_z(\gamma, \alpha) := \lim_{L \rightarrow \infty} m_w(\gamma, \alpha, L) = \int_{t=0}^{\infty} z(t) dt. \quad (22)$$

Figure 9 displays the function  $m_w(\gamma, \alpha, L)$  for  $\gamma = 20$ . Again, periodicity in  $L$  is in accordance with (11), and, interestingly, periodicity holds even for relatively small values of  $L$  (but  $L + \alpha = 1.5$  means  $N \sim 90$  components in this case). The minimum of the mean appears near integer values (which would correspond to  $\alpha = 0$ ); actually, for  $\gamma = 20$ , the mean time to connect has a minimum around  $\alpha \approx 0.97$ .



**Fig. 9.** Mean of the scaled connection time for  $\gamma = 20$



**Fig. 10.** Mean of the scaled connection time for  $\gamma = 2$

Figure 10 displays the function  $m_w(\gamma, \alpha, L)$  for  $\gamma = 2$ . In this case, there is a brief increasing section for smaller values of  $L$  (but  $L = 6$  still corresponds to  $N = 2^6 = 64$  components). In Figure 10, the periodicity in  $L$  is not visible, even though it still occurs, but with a very small amplitude. This smaller amplitude for lower  $\gamma$  is the result of the same phenomenon that was observed in Figures 5 and 6 earlier.

It is also visible that the mean is much larger for  $\gamma = 20$  than for  $\gamma = 2$  (around 15 for  $\gamma = 20$  compared to around 2.7 for  $\gamma = 2$ ). This brings up the question of the optimal value of  $\gamma$ . Due to the facts that

- $m_w(\gamma, \alpha, L) \rightarrow m_z(\gamma, \alpha)$  as  $L \rightarrow \infty$ ,
- $m_w(\gamma, \alpha, L)$  has its minimum near  $\alpha = 0$ , and
- for smaller values of  $\gamma$ ,  $z(\alpha, \gamma, t)$  and thus  $m_z(\gamma, \alpha)$  are near-constant in  $\alpha$  (see also Figure 6),

we examine  $m_w(\gamma, 0, 30)$  for optimization in  $\gamma$ . The result is depicted in Figure 11. It indicates that the optimal value of  $\gamma$  is around 1.65. Interestingly, formula (26) in [10] provides an optimal value of  $\gamma = 1/(1 - 1/e) \approx 1.58$ , albeit for a different setting (stationary with arrivals). That said, in practice,  $\gamma = 2$  might also be a viable option since it is much easier to implement and the mean connection time is not much higher.

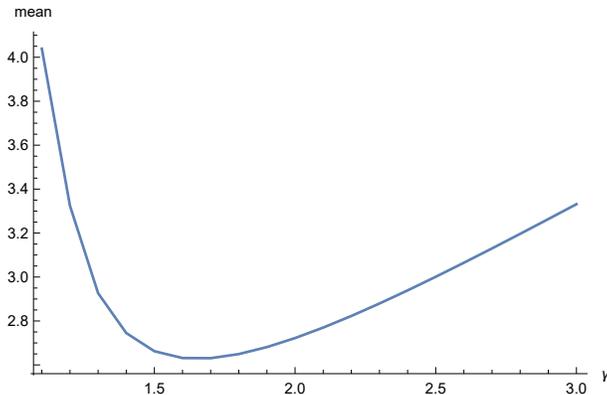


Fig. 11. Mean of the scaled connection time as a function of  $\gamma$

## 5 Additional improvement: switch to no backoff

In this section, we introduce and analyze an additional option that improves the behavior of the system. Since we are interested in further improvement to the results of Section 4, we will focus on smaller values of  $\gamma$ , where the log-periodic behavior is less prominent. Accordingly, the log-periodic behavior will not be emphasized and we will usually settle for optimization for  $\alpha = 0$ , similarly to Figure 11.

In Section 4, we analyzed the mean time to connect. In Remark 7, we concluded that the decay of  $z(t)$  is superexponential. A consequence of this is that the tail of  $z(t)$  has a considerable contribution to the mean time to connect  $m_z$ ; in other words, larger values of  $t$  also contribute to  $m_z$  in the integral (22).

The reason for the heavy tail of  $z(t)$  is due to the backoff policy: while backoff is necessary early on in the process to reduce the probability of collisions, later on, it causes some of the components to “overshoot”: they end up in a relatively high class, and, due to the exponential backoff policy, these components will then take a long time to connect. This effect is more prominent for larger values of  $\gamma$ , but it is significant for any value of  $\gamma$  (see also Figures 5 and 6).

In the literature, this effect is handled by maximizing the number of times a tag performs a backoff, and after reaching the maximum, the tag is dropped (see

EB- $M$  in [11]), resulting in data loss. In order to avoid data loss, we introduce the following, slightly softer approach: the backoff is turned off at some point during the connection process. We will call this “switch to no backoff”. For the original description, this means that the equation for the transition (2) is simply replaced by

$$N_i(n+1) = N_i(n) \quad (23)$$

in case of collision. For the corresponding behaviour for  $z_i(t)$ , we replace (12) by

$$\frac{dz_i(t)}{dt} = -\exp\left(-\sum_{j=-\infty}^{\infty} \gamma^\alpha \gamma^{-j} z_j(t)\right) \gamma^\alpha \gamma^{-(i-1)} z_{i-1}(t). \quad (24)$$

Overall, this means that for some “switching time”  $n_0$ , the behavior of  $N_i(n)$  is governed by (1) and (2) for  $n \leq n_0$  and by (1) and (23) for  $n > n_0$ ; and, correspondingly, the behaviour of  $z_i(t)$  is governed by (12) for  $t \leq t_0$  and by (24) for  $t \geq t_0$ . The corresponding random process will be denoted by  $\bar{N}_i(n)$  and the solution of the differential equation by  $\bar{z}_i(t)$ . (The connection between  $t_0$  and  $n_0$  is simply  $n_0 = t_0 N$  in accordance with the time scaling.)

In practice, switching can be implemented by the channel server sending a signal to all unconnected components to switch to no backoff; components will then simply keep trying to access the channel with the probability  $p_i$  locked in for each component. The main issue is to determine the optimal switching time. For a simulation of  $\bar{N}_i(n)$  or a numerical solution for  $\bar{z}_i(t)$ , we may set the switching time explicitly, but in actual real life application, a global time scale of the process is not available explicitly for the channel server, so the value of  $n_0$  (or  $t_0$ ) can not be set directly.

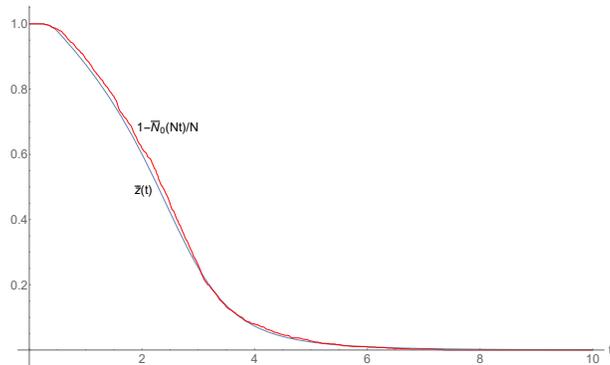
Instead of setting  $t_0$  directly, we will look at the connection rate  $\lambda(t) \exp(-\lambda(t))$  (see (16)). Recalling from Remark 5 that initially,  $\lambda(t)$  is decreasing and  $\lambda(t) \exp(-\lambda(t))$  is increasing before it hits its maximal value  $e^{-1}$ , it follows that there is a bijection between  $t$  and  $\lambda(t) \exp(-\lambda(t))$  before the maximum. The bijection is given by the deterministic relation (16).

So instead of setting the value of  $t_0$  directly, we set the value of the successful connection rate  $\lambda(t_0) \exp(-\lambda(t_0))$ . This can be done in the actual system by keeping track of the number of successful connections in a given time window, and switching when the ratio of successful connections reaches a given threshold. The time when the threshold  $\lambda(t_0) \exp(-\lambda(t_0))$  is reached may differ slightly from the actual value of  $t_0$  due to randomness. A large window allows for a more precise estimation of  $\lambda(t_0) \exp(-\lambda(t_0))$  (and  $t_0$ ), but it also means more bookkeeping for the channel server and is only viable when the component size is very large. We do not analyze the effect of the window size any further.

Again, we compare simulation of  $\bar{N}_i(n)$  with its mean-field counterpart  $\bar{z}_i(t)$ , that is, we check the validity of

$$\bar{z}(t) \approx 1 - \bar{N}_0(n)/N \quad (25)$$

where  $n = tN$  and the switch occurring at  $n_0 = t_0N$ . Figure 12 displays the simulation result for  $N = 1024$  and switching time  $t_0 = 0.5$ . The matching is very good already for  $N = 1024$ . This allows us to examine the behavior of  $\bar{z}(t)$  and make conclusions for both  $z(t)$  and  $N_0(n)$ , similarly to Section 3.2.



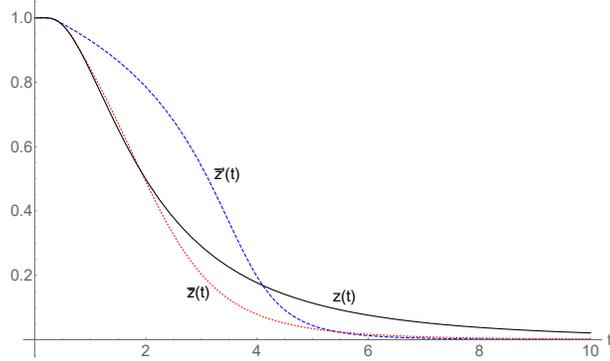
**Fig. 12.** Simulation for  $1 - \bar{N}_0(Nt)/N$  (red line) versus  $\bar{z}(t)$  (dashed blue line); parameters are  $N = 2^{10}$ ,  $\gamma = 2$ ,  $L = 10$ ,  $\alpha = 0$ ,  $t_0 = 0.5$

Figure 13 shows the effect of switching by displaying  $\bar{z}(t)$  (switching at  $t_0 = 0.72$ , dotted red line) versus  $\bar{z}'(t)$  (switching at  $t_0 = 0.39$ , dashed blue line) versus  $z(t)$  (no switching, red line). Switching affects both the bulk and the tail of the connection process. Immediately after the switch, the connection process is slower (compared to no switching), but it accelerates later on, decreasing the tail of the process considerably. Overall, switching offers a tradeoff displayed prominently in Figure 13: an earlier switch results in a prolonged period of slower connection for the bulk, but offers a faster connection for the tail. In Figure 13, eventually  $\bar{z}(t)$  (switching at time  $t_0 = 0.39$ , dashed blue line) decays faster than  $\bar{z}'(t)$  (switching at time  $t_0 = 0.72$ , dotted red line).

The tradeoff between the bulk and the tail displayed in Figure 13 means that optimization for the mean connection time and optimization for a given “quantile”, that is, the time when a prescribed percentage of components has already connected may give entirely different results.

Table 1 shows the 90%, 95%, 99% and 99.9% quantiles of the time to connect along with the mean time to connect for various switching times (including  $\infty$ , that is, the original process with no switching) for both  $\gamma = 2$  and  $\gamma = 1.65$ . The switching times included are actually the optimal values for either the mean or for one of the quantiles; the value for which the switching time is optimal is indicated in boldface in each column. We included all these quantiles and the associated mean for all the switching times in Table 1.

Optimization, e.g. for the 90% quantile, means that we find the value of the switching time,  $t_0$ , such that the time when  $\bar{z}(t)$  reaches  $1 - 0.9$  (which



**Fig. 13.**  $z(t)$  (no switching, black line) versus  $\bar{z}(t)$  (switching at time  $t_0 = 0.72$ , optimal for  $m_z$ , dotted red line) versus  $z'(t)$  (switching at time  $t_0 = 0.39$ , optimal for the 99.9% quantile, dashed blue line). Parameters are  $\gamma = 2, L = 10, \alpha = 0$

corresponds to the time when 90% of the components are connected) is the smallest.

$\gamma$	switching time $t_0$	mean time to connect	quantile			
			0.9	0.95	0.99	0.999
2	$\infty$	2.722	5.306	7.171	12.91	25.47
2	0.718	<b>2.198</b>	3.738	4.522	6.791	11.57
2	0.607	2.230	<b>3.687</b>	4.369	6.328	10.44
2	0.534	2.321	3.732	<b>4.344</b>	6.089	9.730
2	0.453	2.561	3.954	4.486	<b>5.983</b>	9.094
2	0.387	3.019	4.448	4.912	6.201	<b>8.877</b>
1.65	$\infty$	2.628	4.746	6.050	9.776	17.20
1.65	1.008	<b>2.321</b>	3.782	4.439	6.213	9.634
1.65	0.838	2.361	<b>3.748</b>	4.313	5.825	8.729
1.65	0.777	2.408	3.775	<b>4.307</b>	5.719	8.428
1.65	0.677	2.563	3.916	4.390	<b>5.637</b>	8.017
1.65	0.573	2.940	4.325	4.737	5.805	<b>7.833</b>

**Table 1.** Optimization of the switching time for a prescribed quantile ( $\alpha = 0$ )

In accordance with our previous remarks, Table 1 shows that for larger quantile, earlier switching time is better.

## 6 Conclusion and outlook

The paper considers a real life positioning system where a large number of tags are trying to connect to the infrastructure with an exponential backoff policy.

Using mean field methodology, we derived a system of deterministic differential equations whose solution approximates the original process and analysed the solution to find that the system exhibits log-periodic behavior in the mean-field limit. Several aspects of the system were identified (like the early rapid transition), and the sensitivity of the system for the backoff exponent was also addressed. The mean field derivation also allowed us to perform numerical optimization for the backoff exponent and introduced and analyzed an additional improvement by allowing the system to switch to no backoff.

The calculations in the paper, while convincing, are nonrigorous, and were validated only by comparing simulation results with numerical solution of the mean-field limit. A mathematically rigorous proof of the mean-field convergence is work in progress.

In the paper, an exponential backoff policy was assumed. Other backoff policies may also be considered (e.g. polynomial decay); in that case, the scaling of time and space might be entirely different, along with other aspects of the system. This is also subject to future research; just for an interesting tease, we mention that some offhand examination indicates that Remark 6 (and specifically (20)) seems to be valid for other backoff policies as well (regardless of the scaling).

## References

1. Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(12):173 – 187, 1999.
2. Grard Ben Arous, Alexander Fribergh, Nina Gantert, and Alan Hammond. Biased random walks on GaltonWatson trees with leaves. *Ann. Probab.*, 40(1):280–338, 01 2012.
3. John Capetanakis. Tree algorithms for packet broadcast channels. *IEEE transactions on information theory*, 25(5):505–515, 1979.
4. Federal Communications Commission et al. Title 47 – Telecommunication: Chapter I – Federal Communications Commission: Subchapter A – General: Part 15 – radio frequency devices. *Federal Communications Commission Regulatory Information*, 2009.
5. Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 2005.
6. International Organization for Standardization. Information technology – Radio frequency identification for item management – Part 6: Parameters for air interface communications at 860 MHz to 960 MHz General. ISO/IEC standard, 2013.
7. C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
8. T. Komorowski, C. Landim, and S. Olla. *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012.
9. Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

10. Byung-Jae Kwak, Nah-Oak Song, and Leonard E. Miller. On the stability of exponential backoff. *Journal of research of the National Institute of Standards and Technology*, 108(4):289–297, 2003.
11. Byung-Jae Kwak, Nah-Oak Song, and Leonard E. Miller. Performance analysis of exponential backoff. *IEEE/ACM Trans. Netw.*, 13(2):343–355, April 2005.
12. Sylvie Méléard and Vincent Bansaye. *Some stochastic models for structured populations: scaling limits and long time behavior*. Springer International Publishing, 2015.
13. Yuan Shen, Henk Wymeersch, and Moe Z. Win. Fundamental limits of wideband localization - part II: Cooperative networks. *Information Theory, IEEE Transactions on*, 56:4981 – 5000, 11 2010.
14. Guowei Shi and Ying Ming. Survey of indoor positioning systems based on ultra-wideband (UWB) technology. In Qing-An Zeng, editor, *Wireless Communications, Networking and Applications*, pages 1269–1278, New Delhi, 2016. Springer India.
15. Mohammad Shurman, Belal Al Shuaib, Mohammad Alsaedeen, Mamoun Al-Mistarihi, and Khalid Darabkh. N-BEB: New backoff algorithm for IEEE 802.11 MAC protocol. 05 2014.
16. IEEE Computer Society. Low-rate wireless personal area networks (lr-wpans). IEEE Standard, 2011.
17. Liyuan Song, Hongliang Zou, and Tingting Zhang. A low complexity asynchronous UWB TDOA localization method. *International Journal of Distributed Sensor Networks*, 11(10):675490, 2015.
18. Dietrich Stauffer and Didier Sornette. Log-periodic oscillations for biased diffusion on random lattice. *Physica A*, 252(3-4):271–277, 1998.