Dipartimento di Informatica
Università del Piemonte Orientale "A. Avogadro"
Via Bellini 25/G, 15100 Alessandria
http://www.di.unipmn.it

Mean Field Methods in Performance Analysis

Authors: Andrea Bobbio[1], *Marco Gribaudo*[2], *Miklos Telek*[3]

[1] *Dipartimento di Informatica, University of Piemonte Orientale, Italy*
[2] *Dipartimento di Informatica, University of Torino, Italy*
[3] *Department of Telecommunications, Technical University of Budapest, Hungary*

bobbio@mfn.unipmn.it, marcog@di.unito.it, telek@webspn.hit.bme.hu

TECHNICAL REPORT TR-INF-2008-03-03-UNIPMN
(March 2008)

Abstract

Modeling and analysing very large stochastic systems composed of interacting entities is a very challenging and complex task. The usual approach, relying on the generation of the whole state space, is bounded by the state space explosion, even if symmetry properties, often included in the model, allow to apply lumping techniques and building the overall model by means of tensor algebra operations.

In this paper we resort to the *mean field* theory. The main idea of the mean field theory is to focus on one particular tagged entity and to replace all interactions with the other entities with an average or effective interaction. The reduction of a multi-body problem into an effective one-body problem makes the solution easier while at the same time taking into account the contribution of an averaged interdependence of the whole system on the specific entity. In this paper, we apply the mean field approach to very large systems of interacting continuous time Markov chains, in which the averaged interaction depends on the distribution of the entity population in each state.

After deriving the mean field main relations for our particular case and after obtaining asymptotic results when the entity population tends to infinity, we report several examples of interacting Markovian queues, showing the potentialities of the proposed technique.

1 Introduction

Complex systems can usually be disaggregated into interacting parts or components where each part can have a local autonomous behavior that depends on the ensemble of the behaviors of the other parts. In recent years, an enormous amount of literature has been devoted to the study of complex systems in biology, economics, social science, physics, computer and communication systems. In this paper, we focus the attention on very large scale stochastic systems, in which the basic entities evolve according to a CTMC, whose infinitesimal generator depends on current state occupied by all the other entities.

The analysis of large scale stochastic systems composed by interacting objects has been mainly faced in the literature by resorting to the superposition of interacting Markov chains or to fluid models. In the first case, the analysis of the system requires the generation of the global state space, defined as the Cartesian product of the state spaces of the CTMC's describing the individual interacting objects. The explosion of the global state space determines the upper bound for the application of the methodology, even if the explosion is usually mitigated by exploiting the symmetry properties often included in the system definition, that allow to apply lumping techniques and to produce the global transition rate matrix by means of tensor algebra operators applied to the local matrices.

Representative attempts in this direction define the interacting objects directly as Markov chains [4, 6], or as finite state automata [12, 13] or as Petri nets [5, 10]. In [12] the local entity is called automaton and the *Stochastic Automata Network (SAN)* is a system composed by interacting automata. In [1], the states of the individual Markov chains are partitioned in classes and the transition rate of each chain depends on the classes of the other chains. A two layer view is also proposed in social networks in [14] where the local level is a chain that depicts an individual player and the global view models the team action as a whole. The compositional approaches are limited by the explosion of the state space.

A particular model of interacting objects for which a set of exact and approximate analysis methods are available is the queueing network model. In this model the objects communicate via customers which visit the network nodes according to some routing rules. We refer to [2] for a recent survey on the related analysis results. In the most common application of queueing

networks the number of objects is finite and the number of states of the objects can be finite and infinite. The case of infinite number of states of finite number of objects can also be approximated with fluid models. Fluid models [9, 7] are able to capture the global behaviour of the system, but they loose the capability of detailing the local behaviour.

In this paper, we focus the attention on very large scale stochastic systems, whose dimensions exceed the capabilities of all the methods based on the generation of the global state space, even if the basic entities evolve according to a CTMC. Especially, we focus on the case when the number of interacting objects grows very large and the number of states of these objects is finite and moderately large. We propose an approximation based on *mean field* method [11, 3]. The *mean field* method focusses on a particular tagged entity, and replaces all the interactions with the other entities with an average interactions. In the present case, each entity is a CTMC described by a local infinitesimal generator whose entries depends on the distribution of the other entities in their state space. In this way, we can model the individuality of each entity, but at the same time its interaction with the whole system. Asymptotic results allow us to consider systems in which the number of entities tends to infinity.

The mean field technique is well known and widely applied in many different areas [11]. The main goal of this paper is to present this methodology in a way which allows its use in the performance evaluation community. By this reason, we put more emphasis on the explanation of the methodology and how the methodology can find application in stochastic modeling rather than in the practical relevance of the considered examples. With this respect [3] had a very similar goal. The main difference is that in [3] the interacting entities are formulated as discrete time Markov models, while in the present paper we take into consideration continuous time Markov models.

Indeed treatment of continuous time models requires different methodology than the one used for discrete time models and we believe that Lemma 2 plays an important role in establishing the relation between these two methodologies.

The paper is organized as follows; Section 2 introduces the mean field idea for interacting CTMC and provides the main theorem and results. Section 3 illustrates a simple example of interacting queues and shows how different dependent strategies for accommodating the incoming customers can be modeled and analysed; the analysis is restricted to identical and indistinguishable entities. Section 4 introduces a new variant, by showing that it is possible to consider entities belonging to different types and provides a possible application example.

2 Mean field method for large CTMC models

There are several efficient methods for constructing and evaluating Markovian models composed by a large finite number of identical interacting entities. The mean field method allows to compute the behaviour of this kind of models when the number of entities tends to infinity and suggests and approximation when the number of entities is large.

Let us assume that we have N identical discrete state entities in the form of CTMC. The state transitions of the CTMCs might depend on the current state of all entities, but cannot depend on the past history of the process. The state of entity ℓ ($\ell = 1, 2, \ldots, N$) at time t is denoted by $X_\ell(t)$

In this Section we assume, as an essential property, that all the entities are identical and indistinguishable. With this assumption, the behaviour of entity i does not depend directly on the particular state of a generic entity j, but it may depend on the global number of entities in each state.

Due to the fact that the entities are identical, the state of a randomly chosen (tagged) entity is denoted by $X(t)$. The state space of each entity, S, is composed by $s = |S|$ states, and $N_i(t)$ denotes the number of entities which are in state i ($\forall i \in S$) at time t. The vector composed by $N_i(t)$ is denoted by $\mathbf{N}(t)$ and by this definition, $\sum_{i=1}^{s} N_i(t) = N$.

The global behavior of the set of N entities forms a CTMC over the state space of size s^N. However, due to the fact that the entities are identical and indistinguishable, the state space can be lumped into the aggregate state space S_L of size $\binom{N+s-1}{s-1}$, where a state of the overall CTMC is identified by the number of entities staying in each state of S, i.e., by $\mathbf{N}(t) = (N_1(t), N_2(t), \ldots, N_s(t))$.

The evolution of the local CTMC is such that there are no synchronous transitions in different entities and the transition rates of a given entity may depend on the global behavior through the actual value of $\mathbf{N}(t)$. With this assumption, the following transition rates govern the evolution of a particular entity

$$\begin{aligned} &K_{ij}(\mathbf{N}(t)) = \\ &\quad \lim_{\Delta \to 0} \frac{1}{\Delta}\ Pr(X(t+\Delta) = j | X(t) = i, \mathbf{N}(t)), \\ &K_{ii}(\mathbf{N}(t)) = -\sum_{j \in S, j \neq i} K_{ij}(\mathbf{N}(t)). \end{aligned} \tag{1}$$

Note that in the condition of (1), $X(t) = i$ means that $\mathbf{N}(t)$ is such that $N_i(t) \geq 1$, since at least the tagged entity is in state i.

To avoid large numbers in $\mathbf{N}(t)$ when N increases, in a way similar to [3], we introduce the normalized vector, $\mathbf{n}(t) = \mathbf{N}(t)/N$, where the entries $\mathbf{n}(t)$, $0 \leq n_i(t) \leq 1$, define the proportion of objects in state i at time t and $\sum_{i \in S} n_i(t) = 1$. Rewriting (1) with the normalized notation results:

$$\begin{aligned} &k_{ij}(\mathbf{n}(t)) = \lim_{\Delta \to 0} \frac{1}{\Delta}\ Pr(X(t+\Delta) = j | X(t) = i, \mathbf{n}(t)), \\ &k_{ii}(\mathbf{n}(t)) = -\sum_{j \in S, j \neq i} k_{ij}(\mathbf{n}(t)). \end{aligned} \tag{2}$$

(1) and (2) describe the same transition matrix, but we use capital letters to denote the integer vector, $\mathbf{N}(t)$, and its related quantities. We use small letters to denote the normalized quantities which are based on $\mathbf{n}(t)$. With this convention, we can rewrite Equation (2) in matrix form

$$\mathbf{k}(\mathbf{n}(t)) = \{k_{ij}(\mathbf{n}(t))\} \tag{3}$$

The mean field method is based on the following essential theorem.

Theorem 1. *The normalized state vector of the lumped process, $\mathbf{n}(t)$, tends to be deterministic as N tends to infinity and satisfies the following differential equation*

$$\frac{d}{dt}\mathbf{n}(t) = \mathbf{n}(t)\ \mathbf{k}(\mathbf{n}(t)) \tag{4}$$

An individual component out of the set s equations (4) can be written as:

$$\frac{d}{dt}n_i(t) = \sum_{j \in S} n_j(t)\ k_{ji}(\mathbf{n}(t)), \tag{5}$$

Proof. We can evaluate the evolution of the $\mathbf{N}(t)$ process, based on the transition matrix, $K_{ij}(\mathbf{N}(t))$, which governs the evolution of all entities. To this end, we investigate the distribution of the random variable $N_i(t+\Delta) - N_i(t)$, when Δ is small enough. We do this in two steps.

i) - First, we investigate the distribution of $N_i(t+\Delta) - N_i(t)$ conditioned to a given value of $N_i(t)$;

ii) - Second, we uncondition according to the distribution of $N_i(t)$, perform the normalization from $\mathbf{N}(t)$ to $\mathbf{n}(t)$ and evaluate the limits as N tends to infinity and Δ tends to zero.

$$
\begin{aligned}
&N_i(t+\Delta) - N_i(t) \mid \mathbf{N}(t) \cong \\
&\begin{cases}
1 & \displaystyle\sum_{j\in S, j\neq i} N_j(t) K_{ji}(\mathbf{N}(t))\Delta + \sigma(\Delta), \\
0 & 1 + N_i(t) K_{ii}(\mathbf{N}(t))\Delta - \\
& \displaystyle\sum_{j\in S, j\neq i} N_j(t) K_{ji}(\mathbf{N}(t))\Delta + \sigma(\Delta), \\
-1 & -N_i(t) K_{ii}(\mathbf{N}(t))\Delta + \sigma(\Delta), \\
> 1 & \sigma(\Delta), \\
< -1 & \sigma(\Delta),
\end{cases}
\end{aligned}
\tag{6}
$$

The lhs of (6) is a random variable, while its rhs contains the potential values and the associated probabilities, in the first and the second column, respectively. The $\cong$ sign indicates that the equation presents only the tangible elements of the distribution.

Lemma 2. *Let Ψ_{ji} $(j \in S, j \neq i)$ be an independent binomially distributed random variables with parameters $N_j(t)$ (number of trials) and $K_{ji}(\mathbf{N}(t))\Delta$ (probability of occurrence) and Φ_i an independent binomially distributed random variable with parameters $N_i(t)$ and $-K_{ii}(\mathbf{N}(t))\Delta$. The distribution of $\Omega = \displaystyle\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i$ is identical with the one in (6).*

A complete proof of the lemma is provided in Appendix A. The main argument of the proof consists in generating two independent binomially distributed random variables with the above characteristics and computing the probability that their sum is equal to 0, 1 or -1 for small Δ.

As a result of Lemma 2 we rewrite (6) as

$$
N_i(t+\Delta) - N_i(t) \mid \mathbf{N}(t) \cong \sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i \tag{7}
$$

Dividing both sides of (6) and (7) by N and using the normalized notation gives

$$
\begin{aligned}
&n_i(t+\Delta) - n_i(t) \mid \mathbf{n}(t) \cong \\
&\begin{cases}
1/N & \displaystyle\sum_{j\in S, j\neq i} N n_j(t)\ k_{ji}\left(\mathbf{n}(t)\right)\Delta + \sigma(\Delta), \\
0 & 1 + N n_i(t) k_{ii}(\mathbf{n}(t))\Delta - \\
& \displaystyle\sum_{j\in S, j\neq i} N n_j(t)\ k_{ji}\left(\mathbf{n}(t)\right)\Delta + \sigma(\Delta), \\
-1/N & -N n_i(t) k_{ii}(\mathbf{n}(t))\Delta + \sigma(\Delta), \\
> 1/N & \sigma(\Delta), \\
< -1/N & \sigma(\Delta),
\end{cases}
\end{aligned}
\tag{8}
$$

and

$$n_i(t+\Delta) - n_i(t) \mid \mathbf{n}(t) \cong \frac{1}{N}\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right) \tag{9}$$

Lemma 3. *When $N \to \infty$ then the rhs of (9) tends to be deterministic*

$$\lim_{N\to\infty} \frac{1}{N}\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right) = \sum_{j\in S} n_j(t) k_{ji}(\mathbf{n}(t))\Delta \ . \tag{10}$$

The proof of the lemma is provided in Appendix B and is based on the limiting behaviour of a Bernoulli random variable according to the strong law of large numbers.

From (10), as N tends to infinity, we obtain

$$n_i(t+\Delta) - n_i(t) \mid \mathbf{n}(t) \cong \sum_{j\in S} n_j(t)\ k_{ji}(\mathbf{n}(t))\Delta, \tag{11}$$

According to (11) $n_i(t)$ remains deterministic for all $t > 0$ if its initial value $n_i(0)$ is deterministic. This way we can write

$$n_i(t+\Delta) - n_i(t) = \sum_{j\in S} n_j(t)\ k_{ji}(\mathbf{n}(t))\Delta + \sigma(\Delta). \tag{12}$$

Dividing both sides of (12) by Δ we obtain (5) as Δ tends to 0.

□

Theorem 1 provides a formulation that is easy to apply and to compute in the extreme case when N tends to infinity. The following corollary investigates the case when the same approach is applied for systems with a finite number of entities.

Corollary 4. *When N is sufficiently large, the normalized state vector of the lumped process, $\mathbf{n}(t)$, is a random vector whose mean can be approximated by the following differential equation*

$$\frac{d}{dt}E(n_i(t)) \approx \sum_{j\in S} E(n_j(t))\ k_{ji}(E(\mathbf{n}(t))) \tag{13}$$

Proof. When N is finite both sides of (9) represent a "real" random variable with positive variance. Taking the expectation of both sides of (9) gives

$$\begin{aligned} & E\left(n_i(t+\Delta) - n_i(t) \mid \mathbf{n}(t)\right) \\ & \cong \frac{1}{N}\ E\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right) \\ & = \sum_{j\in S} n_j(t) k_{ji}(\mathbf{n}(t))\Delta \end{aligned} \tag{14}$$

To obtain an unconditional expression based on (14) we weight the expression according to the distribution of the condition.

$$\begin{aligned} & E\Big(n_i(t+\Delta) - n_i(t)\Big) \\ &= \sum_{\mathbf{m}\in S_L} Pr(\mathbf{n}(t)=\mathbf{m}) \\ &\qquad E\Big(n_i(t+\Delta) - n_i(t) \mid \mathbf{n}(t)=\mathbf{m}\Big) \\ &\cong \sum_{\mathbf{m}\in S^N} Pr(\mathbf{n}(t)=\mathbf{m}) \sum_{j\in S} m_j \; k_{ji}(\mathbf{m})\Delta \end{aligned} \tag{15}$$

If $\mathbf{n}(t)$ is close to deterministic, then the probability that $\mathbf{n}(t)$ is very close to its mean is close to one,

$$\sum_{\mathbf{m}:|E(\mathbf{n}(t))-\mathbf{m}|<\epsilon} Pr(\mathbf{n}(t)=\mathbf{m}) \approx 1,$$

where ϵ is a small vector.

If $\mathbf{n}(t)$ enjoys the property to be close to deterministic, then those terms in which $\mathbf{m}$ is far from $E(\mathbf{n}(t))$ vanishe in the rhs of (15) and we obtain the following approximation

$$\begin{aligned} & E\Big(n_i(t+\Delta) - n_i(t)\Big) \\ &\quad \approx \sum_{j\in S} E(n_j(t)) \; k_{ji}(E(\mathbf{n}(t)))\Delta \; . \end{aligned} \tag{16}$$

Starting from this equation, standard analysis steps result in

$$\begin{aligned} & \frac{E(n_i(t+\Delta)) - E(n_i(t))}{\Delta} \\ &\quad \approx \sum_{j\in S} E(n_j(t)) \; k_{ji}(E(\mathbf{n}(t))) \end{aligned} \tag{17}$$

from which we obtain (13) as Δ tends to 0. □

3 Mean field analysis of dependent queues

To demonstrate the mean field methodology we present a simple example and detail its analysis according to the concepts and quantities discussed in the previous section.

Let us consider a queueing system composed by N identical Markovian queues (entities), each of which has a single server and a buffer of size 1 ($S = \{0, 1, 2\}, s = 3$). Customers arrive at rate $N\lambda$ to this queueing system and their service time is exponentially distributed with parameter μ. The CTMC of a single queue ($N = 1$) is depicted in Figure 1.

3.1 The incoming customer chooses the shortest queue

When $N > 1$ we adopt a policy that the incoming customer chooses the shortest queue and is , thus, directed to the queue which has the less number of customers in it. This policy makes the different queues interdependent. When $N = 2$, the CTMC describing this behavior is depicted in Figure 2, where the first number refers to the state of queue 1 and the second to the state

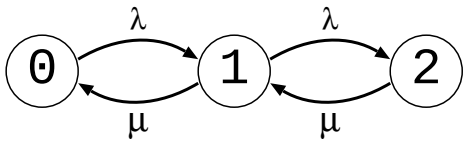

Figure 1: Markov chain of a single queue in isolation

of queue 2. We can interpret the transitions of Figure 2 from the view point of queue 1. In this case, the transition rates of the arrivals to queue 1 depend on the state of queue 2 as it is depicted in Figure 3.

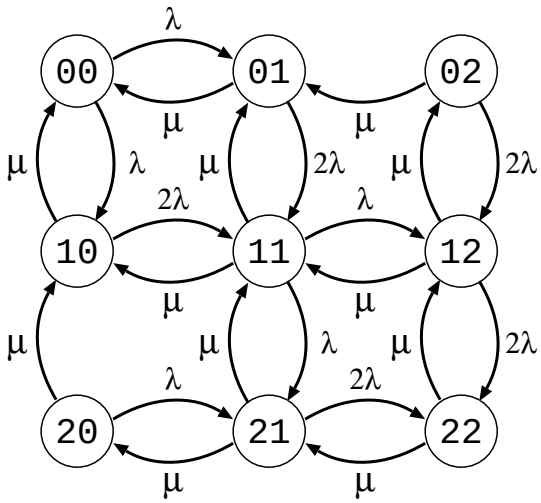

Figure 2: Markov chain of 2 queues (without lumping)

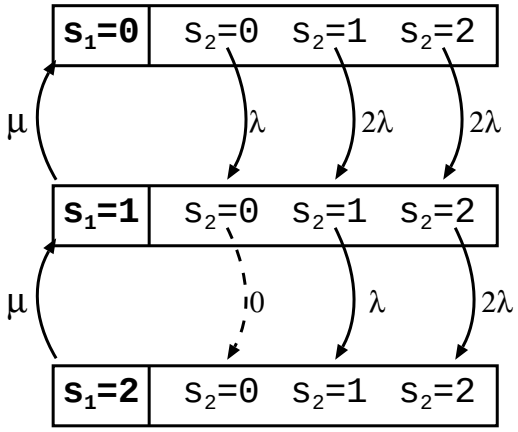

Figure 3: Dependence of the transitions of queue 1 on the state of queue 2

Since the queues are identical in our model we can lump the states according to Figure 4 and we obtain the CTMC depicted in Figure 5. The lumped state space is composed by $\binom{2+3-1}{3-1} = 6$ states, $\mathbf{N}(t) \in \{(2,0,0),(1,1,0),(0,2,0),(1,0,1),(0,0,2),(0,1,1)\}$, where the states are identified by the number of queues having a given number of customers in it. E.g., state $(1,1,0)$ means that one of the queues is idle and one of them has 1 customer in it.

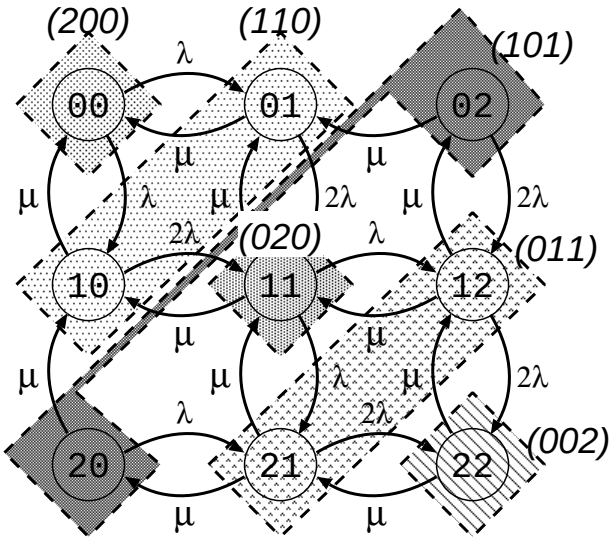

Figure 4: Lumping the Markov chain of 2 queues

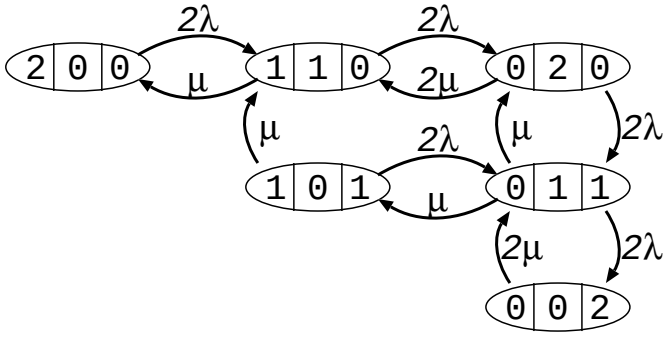

Figure 5: Markov chain of the overall behaviour

Considering the lumped process, we can interpret the behavior from the point of view of a tagged queue. In this case the arrival rates depend on the states of the lumped CTMC, as it is in Figure 6

Due to the fact that all queues are identical Figure 6 contains all information about the process. Consequently, Figure 6 and Figure 5 give an equivalent description of the process. To keep the system description compact (and independent of N) the description of a single tagged entity (the one in Figure 6) is used in practice.

For example, our queueing system can be described as the Markov chain in Figure 7, where

$$\Lambda_0(\mathbf{N}(t)) = \begin{cases} \lambda & \text{if } \mathbf{N}(t) = (2,0,0), \\ 2\lambda & \text{if } \mathbf{N}(t) = (1,1,0), \end{cases}$$

$$\Lambda_1(\mathbf{N}(t)) = \begin{cases} 0 & \text{if } \mathbf{N}(t) = (1,1,0), \\ \lambda & \text{if } \mathbf{N}(t) = (0,2,0), \\ 2\lambda & \text{if } \mathbf{N}(t) = (0,1,1). \end{cases}$$

Having this compact system description, the only remaining step is to introduce the normalized occupancy vector $\mathbf{n}(t) = \mathbf{N}(t)/N$. Doing this, we get

$$\mathbf{n}(t) \in \{(1,0,0), (0.5,0.5,0), (0,1,0), \\ (0.5,0,0.5), (0,0,1), (0,0.5,0.5)\}$$

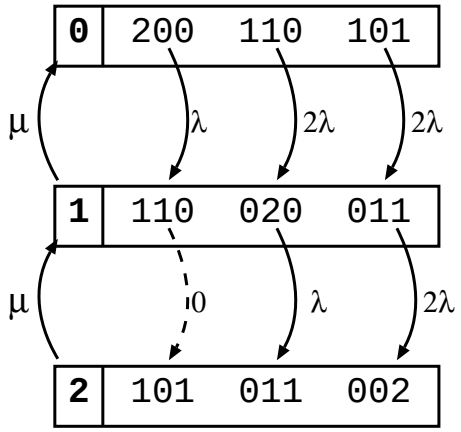

Figure 6: Dependence of the transitions of the tagged queue on the lumped state

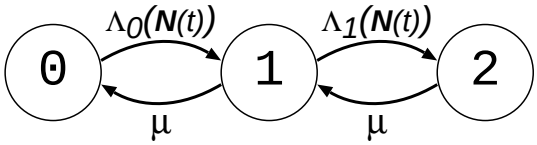

Figure 7: Markov chain of one of the identical entities

and

$$\lambda_0(\mathbf{n}(t)) = \begin{cases} \lambda & \text{if } \mathbf{n}(t) = (1,0,0), \\ 2\lambda & \text{if } \mathbf{n}(t) = (0.5,0.5,0), \end{cases}$$
$$\lambda_1(\mathbf{n}(t)) = \begin{cases} 0 & \text{if } \mathbf{n}(t) = (0.5,0.5,0), \\ \lambda & \text{if } \mathbf{n}(t) = (0,1,0), \\ 2\lambda & \text{if } \mathbf{n}(t) = (0,0.5,0.5). \end{cases}$$

To make the system description independent of N we can rewrite the transition rates as

$$\lambda_0(\mathbf{n}(t)) = \frac{\lambda}{n_0(t)},$$
$$\lambda_1(\mathbf{n}(t)) = \begin{cases} 0 & \text{if } n_0(t) > 0, \\ \dfrac{\lambda}{n_1(t)} & \text{if } n_0(t) = 0, \end{cases} \tag{18}$$

and obtain the CTMC shown in Figure 8. Note that the transitions with fixed rate are the transitions which are independent of the state of the other entities, while the transitions that are function of the occupancy vector represent the dependency of the entities, that is the action of the mean field.

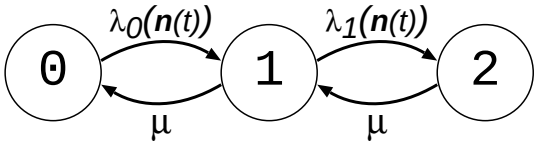

Figure 8: Markov chain of one of the identical entities

The N independent description of the system in (18) is the key to evaluate the limiting

behaviour when N tends to ∞. In this case, the particular form of (5) is

$$\begin{aligned}
&\frac{d}{dt}\{n_0(t), n_1(t), n_2(t)\} \\
&= \{n_0(t), n_1(t), n_2(t)\} \\
&\quad \begin{bmatrix} -\lambda_0(\mathbf{n}(t)) & \lambda_0(\mathbf{n}(t)) & 0 \\ \mu & -\mu - \lambda_1(\mathbf{n}(t)) & \lambda_1(\mathbf{n}(t)) \\ 0 & \mu & -\mu \end{bmatrix}
\end{aligned} \tag{19}$$

and starting from $\mathbf{n}(0) = \{1, 0, 0\}$ this differential equation results in

$$\lim_{t\to\infty} \mathbf{n}(t) = \begin{cases} \{1 - \frac{\lambda}{\mu}, \frac{\lambda}{\mu}, 0\} & \text{if } \frac{\lambda}{\mu} < 1, \\ \{0, 0, 1\} & \text{if } 1 \leq \frac{\lambda}{\mu}. \end{cases} \tag{20}$$

Note that $\lambda_0(\mathbf{n}(t)) = \dfrac{\lambda}{n_0(t)}$ is always multiplied by $n_0(t)$ in the rhs of (19), hence the limit when $n_0(t)$ tends to 0 does not cause problem in the solution. The same situation occurs with $\lambda_1(\mathbf{n}(t))$ and $n_1(t)$.

Figure 9 shows how the limit proposed in equation (20) actually holds, by plotting the mean number of entities in each state as a function of N, where $\lambda = 1.5$ and $\mu = 2$.

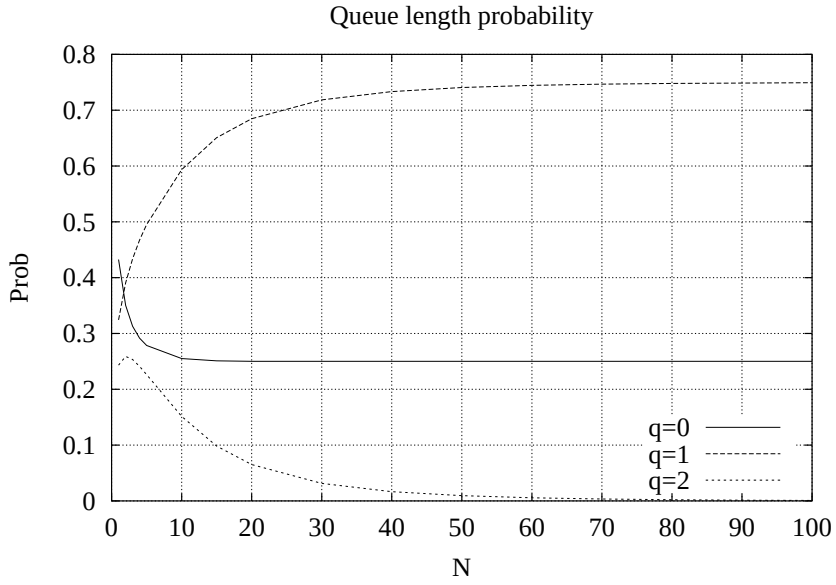

Figure 9: Queue length probability as function of N

3.2 Convergence when N tends to infinity

In order to demonstrate the convergence to a deterministic quantity of the occupancy vector $\mathbf{n}(t)$ when N tends to infinity, we have considered a system of N entities where each entity is limited to 2 states (that is, each queue can only be empty or in service).

In this way the complete stationary occupancy vector, $\{n_0, n_1\}$, can be univocally defined by a single random number, n_0. The results of the exact computations over the complete system are reported in Figures 10 and 11, when $\lambda = 1.5$ and $\mu = 2$.

Figure 10 shows how the coefficient of variation of the number of queues in the first state tends to zero as N tends to infinity in both linear and log-log scale. The log-log plot makes it evident that the decreasing behaviour of the coefficient of variation has a slope proportional to $\sqrt{N}$ as stated by the strong law of large numbers. Figure 11 plots instead the whole distribution of the number of entities in the first state, and shows how it tends to become deterministic.

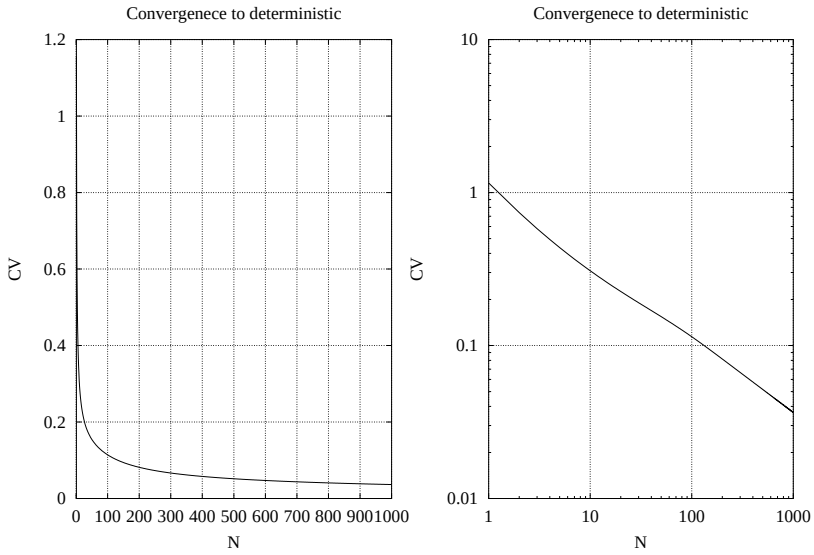

Figure 10: Cv of the exact system as function of N (linear and logarithmic scale)

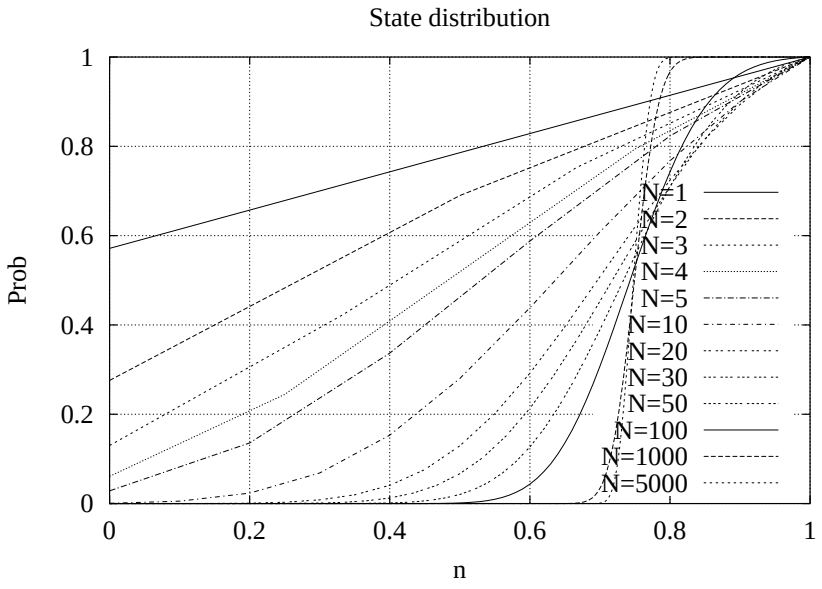

Figure 11: Distribution of the exact system as function of N

3.3 The incoming customer chooses the shortest of K queues

Other management policies, introducing different and more complex dependencies among the queues, are also easy to model and analyze with the mean field method. A variant of the

previous example is when the new incoming customer randomly selects K queues and it joins the one with the less customers out of the selected K queues. This variant represents the random queue selection (independent case) when $K = 1$ and it represents the shortest queue selection of Section 3.1 when $K = N$. The subsequent analysis assumes a fix K independent of N.

According to this policy, the probability that an arriving customer attends a queue with i customers in it can be computed as follows

$$\begin{array}{l} Pr(\text{new customer goes to queue of length } i) = \\ Pr(K \text{ selected queues are longer than } i-1 \\ \quad \text{and at least one selected queue has length } i) = \\ Pr(K \text{ selected queues are longer than } i-1) - \\ \quad Pr(K \text{ selected queues are longer than } i) \end{array} \tag{21}$$

To compute these probabilities we introduce the following notation. The number of queues with at least i customers in it is $S_i(t) = \sum_{j=i}^{s} N_j(t)$. The proportion of queues with at least i customers in it is $s_i(t) = S_i(t)/N = \sum_{j=i}^{s} n_j(t)$. Using these notations

$$\begin{array}{l} Pr(K \text{ selected queues are longer than } i-1) = \\ \dfrac{S_i(t)}{N} \cdot \dfrac{S_i(t)-1}{N-1} \cdots \dfrac{S_i(t)-K+1}{N-K+1} = \dfrac{\binom{S_i(t)}{K}}{\binom{N}{K}} \; . \end{array} \tag{22}$$

When N tends to infinity we have

$$\begin{array}{l} \lim\limits_{N\to\infty} Pr(K \text{ selected queues are longer than } i-1) \\ \quad = s_i(t)^K \; . \end{array} \tag{23}$$

Based on (22), the overall arrival rate towards the queues of length i is $\lambda N \dfrac{\binom{S_i(t)}{K} - \binom{S_{i+1}(t)}{K}}{\binom{N}{K}}$

and the arrival rate to one of the queues of length i is

$$\Lambda_i(\mathbf{N}(t)) = \frac{\lambda N}{N_i(t)} \cdot \frac{\binom{S_i(t)}{K} - \binom{S_{i+1}(t)}{K}}{\binom{N}{K}} \; . \tag{24}$$

Similarly when N tends to infinity, from (23), we have

$$\lambda_i(\mathbf{n}(t)) = \frac{\lambda}{n_i(t)} \left(s_i(t)^K - s_{i+1}(t)^K \right) \; . \tag{25}$$

We have implemented the mean field analysis of the above detailed queue selection policy when each queue has at most 3 customers and we have evaluated the system behaviour in two cases:

Case i) - light load, $\rho = \lambda/\mu = 0.5$ $(\lambda = 1, \mu = 2)$

Case ii) - heavy load, $\rho = \lambda/\mu = 2$ ($\lambda = 2, \mu = 1$).

As a result of the mean field analysis, we have depicted in Figures 12 and 13 the mean queue length for the light and heavy load, respectively.

We observe different trends in the light and the heavy loaded cases. Under light load (Figure 12), the selection of the shortest queue ($K = N$) means that half of the queues have 1 customer and half of them are idle. Instead, with a random queue selection ($K = 1$) the probability of having some queues with 2 customers is positive and the mean queue length is higher.

In case of heavy load (Figure 13), the selection of the shortest queue ($K = N$) means that all the queues are going to be saturated (i.e., in state 2) with probability 1. Instead in case of random queue selection ($K = 1$) the probability that a significant portion of the queues is not selected for a long time is so high that the probability of having less than 2 customers in a significant portion of the queues is positive. As a result the mean queue length is less in this case.

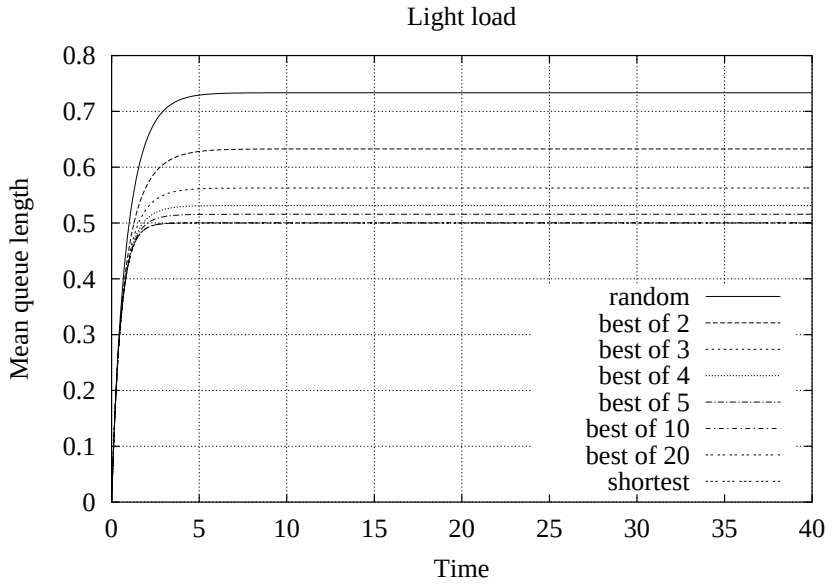

Figure 12: System behaviour with light load

Another important performance metric that can be computed from the model is the mean loss probability of an incoming customer, that is the probability that a client is routed to a queue which is already full and cannot hold it. This metric is computed as follows:

$$Pr(loss) = \frac{E(\#\text{ incoming customers}) \; - \; E(\#\text{ served customers})}{E(\#\text{ incoming customers})} = \lim_{t\to\infty} \frac{N\lambda - (N - N_0(t))\mu}{N\lambda} = \lim_{t\to\infty} \frac{\lambda - (1 - n_0(t))\mu}{\lambda} \tag{26}$$

Figure 14 shows this quantity for both the light and the heavy loaded cases as a function of K. In both cases, increasing K reduces the loss probability of the system. When the system is light loaded (i.e. $\lambda < \mu$), it is sufficient to have $K \geq 4$ to obtain loss probabilities smaller than the machine precision. For the heavy loaded case (i.e. $\lambda > \mu$), the probability does not tend to 0, but to $\frac{\lambda-\mu}{\lambda}$. In order to better understand how the loss probability reaches this limit, in

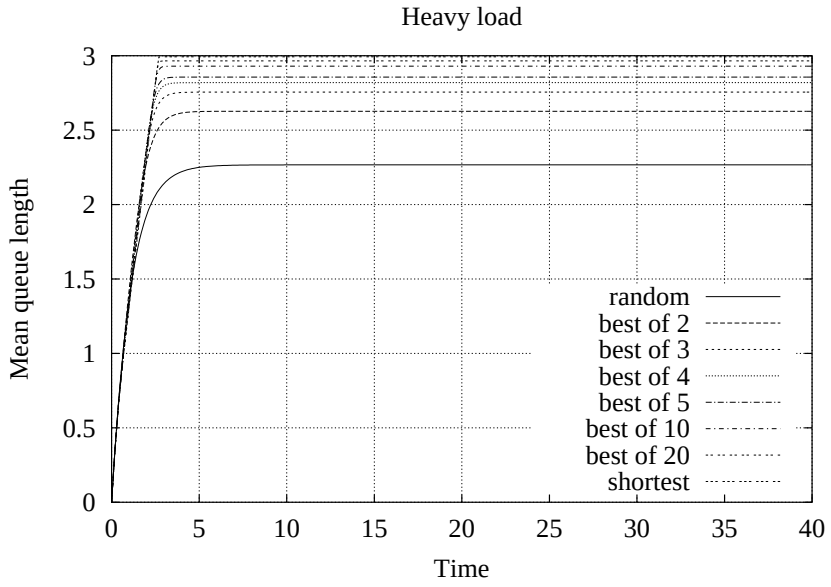

Figure 13: System behaviour with heavy load

the logarithmic version of Figure 14, we have plotted $Pr(loss) - \frac{\lambda-\mu}{\lambda}$. As it can be seen, when the system is heavily loaded, K must be increased much more than in the lightly loaded case to reduce the losses.

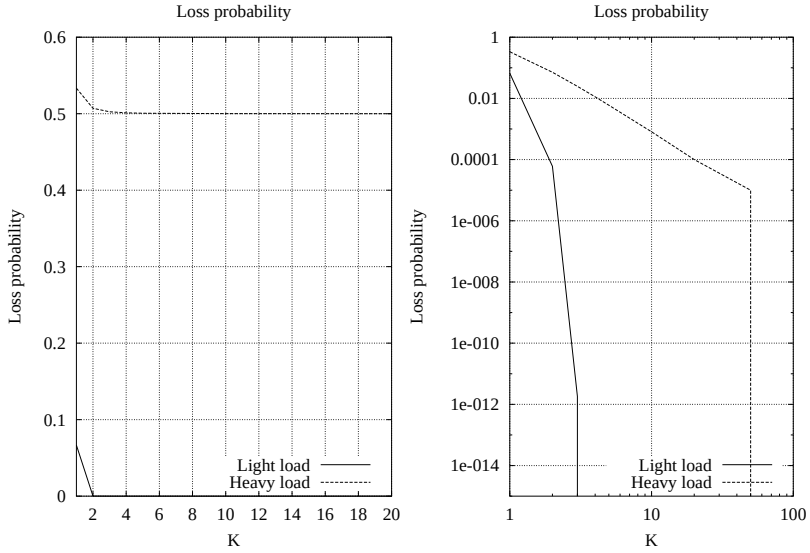

Figure 14: Loss probability (linear and logarithmic scale)

3.4 Comparison with finite N systems

In practice systems are finite, and the assumption of N that tends to the infinity is often non-realistic. One of the questions is thus whether the mean field approach is appropriate to approximate finite systems, and up to which extent. In Corollary 4 we have already proven that this approximation is valid for $N \to \infty$. In Figure 15, we elaborate on this problem, focusing on the loss probability of the heavy loaded model of Section 3.3, and comparing the

loss performance index on a finite system with N queues, using discrete event simulation, with the results obtained from the mean field analysis with $N \to \infty$. The solid line in Figure 15 refers to the mean field computation with $N \to \infty$ and $K = 3$ and the dotted line to the mean field analysis with $K = 2$. It can be seen, that the 95% confidence intervals computed by simulation with $N = 2;\ K = 2$ and with $N = 3;\ K = 3$ are well centered on the asymptotic mean field results. Moreover, for $N \geq 10$ the simulated confidence intervals of the loss probability shrink around the asymptotic mean field values.

The above results indicate that the solutions computed via mean field analysis, can be considered as a meaningful approximation of the exact performance indices, even for relatively small (i.e. $N = 10$) numbers of entities in the system.

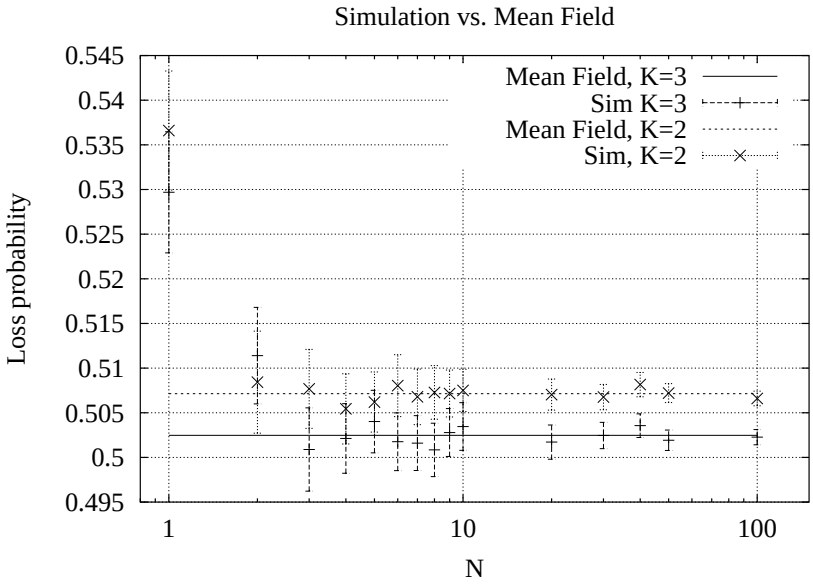

Figure 15: Comparison of the loss probability between mean field, and simulation of a finite system with N queues

4 Mean field method with different kinds of entities

In section 2, we considered the mean field analysis of N dependent identical Markovian entities. In this section we extend the analysis to systems composed by more than one type of dependent Markovian entities. Let $N^{(1)}, N^{(2)}, \ldots, N^{(C)}$ be the number of identical entities of type $1, 2, \ldots, C$, respectively. The state space of a type c entity is denoted by $S^{(c)}$ ($c \in \{1, 2, \ldots, C\}$), and is composed by $s^{(c)} = |S^{(c)}|$ states. $N_i^{(c)}(t)$ denotes the number of type c entities which are in state i at time t. We introduce vector $\mathbf{N}^{(c)}(t)$ of size $s^{(c)}$, whose elements are $N_i^{(c)}(t)$ and vector $\mathbf{N}(t)$ of size $s = \sum_{c=1}^{C} s^{(c)}$, whose blocks are $\mathbf{N}^{(c)}(t)$.

In general, the transition matrix of the Markov chain of a type c entity may depend on the whole vector $\mathbf{N}(t)$. The transition rates of a type c entity are

$$K_{ij}^{(c)}(\mathbf{N}(t)) =$$

$$\lim_{\Delta \to 0} \frac{1}{\Delta}\ Pr(X^{(c)}(t+\Delta) = j | X^{(c)}(t) = i, \mathbf{N}(t)),$$

$$K_{ii}^{(c)}(\mathbf{N}(t)) = -\sum_{j \in S, j \neq i} K_{ij}^{(c)}(\mathbf{N}(t)).$$

Introducing again, $\mathbf{n}(t) = \mathbf{N}(t)/N$ and taking the limit $N \to \infty$ such that $N^{(c)}/N$ remains constant [1] we obtain the same differential equation as (5), but this time vector $\mathbf{n}(t)$, contains the proportion of entities of each type in each state.

As a result, the cardinality of the differential equation (5) increases linearly with the number of different types, since matrix $\mathbf{k}(\mathbf{n}(t))$ is of size $s \times s$ (with $s = \sum_{c=1}^{C} s^{(c)}$) and is composed by non-zero blocks of size $s^{(c)} \times s^{(c)}$.

4.1 Example of a system with regular and spare queues

Let us consider the queueing system of Section 3 with two types of queues: regular queues and *spare* queues. For each regular queue, there are γ spares. For example $\gamma = 0.5$ means that there is a spare every 2 regular queues. Both regular and spare queues have the same service rate μ, and the same buffer capacity B. Costumers arrive at rate λ per regular queue, and are directed to the queue with the lowest occupancy. Spare queues are used only if the mean number of costumers in the regular queues exceeds a given threshold β. We call $\alpha = \frac{1}{1+\gamma}$ the fraction of regular queues. Since the arrival rate is expressed *per regular queue*, we compute the total arrival rate as $\tilde{\lambda} = \alpha\lambda$.

We can apply mean field analysis to this system considering regular queues (identified by vector $\mathbf{n}^{(R)}$), and spare queues (identified by vector $\mathbf{n}^{(S)}$) separately. In particular, if we consider $B = 1$ for sake of simplicity, we have:

$$\begin{aligned} \mathbf{n}(t) &= \{n_0^{(R)}(t), n_1^{(R)}(t), n_0^{(S)}(t), n_1^{(S)}(t)\} \\ \mathbf{n}(0) &= \{\alpha, 0, 1-\alpha, 0\} \end{aligned}$$

$$\mathbf{k}(\mathbf{n}(t)) = \begin{bmatrix} -\lambda_R(\mathbf{n}(t)) & \lambda_R(\mathbf{n}(t)) & 0 & 0 \\ \mu & -\mu & 0 & 0 \\ 0 & 0 & -\lambda_S(\mathbf{n}(t)) & \lambda_S(\mathbf{n}(t)) \\ 0 & 0 & \mu & -\mu \end{bmatrix}$$

where:

$$\lambda_R = \begin{cases} \dfrac{\tilde{\lambda}}{n_0^{(R)}(t)} & \text{if } n_1^{(R)}(t) \leq \alpha\beta \\ \dfrac{\tilde{\lambda}}{n_0^{(R)}(t) + n_0^{(S)}(t)} & \text{if } n_1^{(R)}(t) > \alpha\beta \end{cases}$$

$$\lambda_S = \begin{cases} 0 & \text{if } n_1^{(R)}(t) \leq \alpha\beta \\ \dfrac{\tilde{\lambda}}{n_0^{(R)}(t) + n_0^{(S)}(t)} & \text{if } n_1^{(R)}(t) > \alpha\beta \end{cases}$$

Figure 16 shows some results for $\gamma = 0.5$, varying both the load of the system $\rho = \frac{\lambda}{\mu}$, and the switching point β. The introduction of the spare queues allows the system to respond to load greater than 1 (up to $\rho = 1 + \gamma$). Spare queues are used only if the total load produces a mean queue length larger than β. After this threshold, the mean queue length of the regular queues remains constant, until the it is reached by the mean length of the spare queues. From

[1] Indeed this condition can be relaxed, it is motivated only by typical behaviour of the applications.

this point on both regular and spare queues grow with the same slope. If we consider a fixed load $\rho < 1$, we can see that large values of β reduce spare queues usage, while small values of β improve the response time by allowing shorter queues.

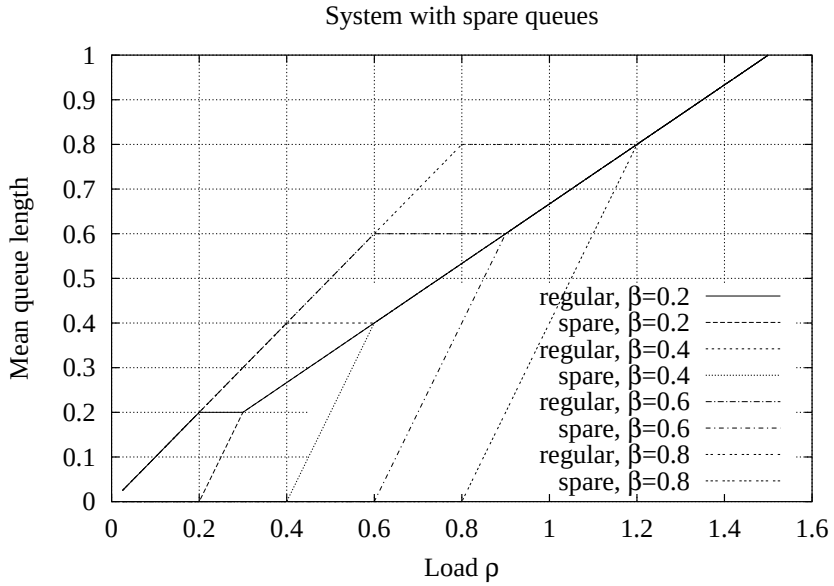

Figure 16: Mean queue length in a system with 50% spare queues

5 Mean field analysis of queues with memory dependent load

Let us consider the previous system queues with two types of customers. The pure load dependent queue selection, in section 4.1, resulted in that the traffic served by the two type of queues was very unbalanced. E.g., asymptotically all customers are directed to the type 1 queues, when $\lambda/\mu < \alpha$. If the goal is to set the traffic served by the different type of queues to a predefined value, β, we can apply a reward based queue selection policy. To this end we define two memory (reward) variables $m_1(t)$ and $m_2(t)$ accumulating the amount of traffic served by type 1 and 2 queues. The associated service (reward) rates are

$$\frac{d}{dt}\, m_1(t) = \sum_{i=0}^{B^{(1)}} r_i^{(1)}\; n_i^{(1)}(t) = \sum_{i=1}^{B^{(1)}} \mu\; n_i^{(1)}(t)\ .$$

I.e., $r_i^{(1)}$ is μ if $i \geq 1$. The reward rates associated with $m_2(t)$ are defined similarly.

An incoming customer is directed to that type of queues which served less customers than the predefined ratio and among the queues of this type it attends the shortest one. I.e., a customer attends a type 1 queue, if $\frac{m_1(t)}{m_2(t)} < \frac{\beta}{1-\beta}$, a type 2 queue if if $\frac{m_1(t)}{m_2(t)} > \frac{\beta}{1-\beta}$, and attends the shortest of all queues if $\frac{m_1(t)}{m_2(t)} = \frac{\beta}{1-\beta}$.

The structure of the transition matrix of type 1 and 2 entities remains the same as the one

in (19) with $B^{(1)}$ and $B^{(2)}$ customers, but the arrival rates depend also on the memory variables

$$\lambda_i^{(1)}(\mathbf{n}(t),\mathbf{m}(t)) = \begin{cases} 0 & \text{if } \frac{m_1(t)}{m_2(t)} > \frac{\beta}{1-\beta} \text{ or} \\ & \left(\frac{m_1(t)}{m_2(t)} = \frac{\beta}{1-\beta} \text{ and} \right. \\ & \left. \quad \exists k<i \text{ s.t. } n_k^{(1)}(t)>0\right), \\ \dfrac{\lambda}{n_i^{(1)}(t)} & \text{if } \frac{m_1(t)}{m_2(t)} < \frac{\beta}{1-\beta}, \\ \dfrac{\lambda}{n_i^{(1)}(t)+n_i^{(2)}(t)} & \text{if } \left(\frac{m_1(t)}{m_2(t)} = \frac{\beta}{1-\beta} \text{ and} \right. \\ & \left. \quad \forall k<i, n_k^{(1)}(t) = n_k^{(2)}(t) = 0\right), \end{cases} \tag{27}$$

The arrival rates to type 2 entities are symmetric with (27).

Assuming that $\frac{\alpha}{\beta} < \frac{1-\alpha}{1-\beta}$ (that is, type 1 gets saturated first), and starting from $\mathbf{n}^{(1)}(0) = \{\alpha, 0, \ldots, 0\}$ and $\mathbf{n}^{(2)}(0) = \{1-\alpha, 0, \ldots, 0\}$ this memory based queue selection policy results in

$$\lim_{t\to\infty} \mathbf{n}^{(1)}(t) = \begin{cases} \{\alpha - \beta\frac{\lambda}{\mu}, \beta\frac{\lambda}{\mu}, 0, \ldots, 0\} & \text{if } \frac{\lambda}{\mu} < \frac{\alpha}{\beta}, \\ \{0, 0, \ldots, 0, \alpha\} & \text{if } \frac{\alpha}{\beta} \le \frac{\lambda}{\mu}, \end{cases}$$

and

$$\lim_{t\to\infty} \mathbf{n}^{(2)}(t) = \begin{cases} \{1-\alpha-(1-\beta)\frac{\lambda}{\mu}, (1-\beta)\frac{\lambda}{\mu}, 0, \ldots, 0\} & \text{if } \frac{\lambda}{\mu} < \frac{\alpha}{\beta}, \\ \{1-\alpha, 0, 0, \ldots, 0\} & \text{if } \frac{\alpha}{\beta} \le \frac{\lambda}{\mu}. \end{cases}$$

6 Conclusion

Mean field theory aims at representing a multi-body problem constituted by a large number of interacting entities with a one-body problem where the interdependencies are replaced by an averaged effective interaction. We have shown how the mean field method can be applied to performance problems where the interacting entities are represented by Continuous Time Markov Chains (CTMC). Asymptotic results have been derived when the number of entities tends to infinity, while approximate results are available with a finite number of entities. We have applied the above method to study different dependent policies for feeding Markovian queues with a finite buffer. The mutual interaction is modeled by defining the transition rates of a tagged customer as a function of the proportion of queues in each state and solving a differential equation defined over the normalized state occupancies. Various performance indices are computed and the behaviour of the interdependent policies is compared with the independent case.

Acknowledgments

The research presented in this paper was partially supported by the EU project CRUTIAL ((http://crutial.cesiricerca.it/) and partially by OTKA grant no. K61709.

Appendix

A Proof of Lemma 2

Proof. If Υ is a binomially distributed random variable with parameters n and $c\Delta$ then

$$Pr(\Upsilon = k) = \binom{n}{k}(c\Delta)^k(1-c\Delta)^{n-k},$$

and

$$\begin{aligned} Pr(\Upsilon = 0) &= (1-c\Delta)^n = 1 - nc\Delta + \sigma(\Delta), \\ Pr(\Upsilon = 1) &= n(c\Delta)^1(1-c\Delta)^{n-1} = nc\Delta + \sigma(\Delta), \\ Pr(\Upsilon = 2) &= \binom{n}{2}(c\Delta)^2(1-c\Delta)^{n-2} = \sigma(\Delta)\ , \end{aligned}$$

because all terms containing Δ^k, $k \geq 2$ belongs to the $\sigma(\Delta)$ class.

No we evaluate the low order Δ terms in the sum of s independent binomially distributed random variables, each of which follows the same pattern as Υ with respect to its dependence on Δ.

The probability of the case when any of the s random variables is equal to or greater than 2 is $\sigma(\Delta)$. The probability of the case when more than one random variables is equal to 1 is also $\sigma(\Delta)$. Hence the dominant cases with low order Δ terms are when each binomially distributed random variable equals to 0 and when one of them equals to 1 and all others equal to 0.

The probability of the case when each binomially distributed random variable equals to 0 is

$$\begin{aligned} &\prod_{j\in S, j\neq i} Pr(\Psi_{ji} = 0) \ \cdot \ Pr(\Phi_i = 0) \\ &= \prod_{j\in S, j\neq i} \Big(1 - N_j(t)K_{ji}(\mathbf{N}(t))\Delta + \sigma(\Delta)\Big) \\ &\quad \cdot \Big(1 + N_i(t)K_{ii}(\mathbf{N}(t))\Delta + \sigma(\Delta)\Big) \\ &= 1 - \sum_{j\in S, j\neq i} N_j(t)K_{ji}(\mathbf{N}(t))\Delta \\ &\quad + N_i(t)K_{ii}(\mathbf{N}(t))\Delta + \sigma(\Delta). \end{aligned}$$

In this case $\Omega = 0$. Ω can be 0 also in other cases, e.g., when Ψ_k and Φ_i equal to 1 and all other Ψ_{ji} equal to 0, but the probability of the possible other cases resulting $\Omega = 0$ is $\sigma(\Delta)$.

The probability of the case when Ψ_{ki} equals to 1 and all other Ψ_{ji} and Φ_i equal to 0 is

$$\begin{aligned} &Pr(\Psi_{ki} = 1) \ \cdot \prod_{j\in S, j\neq i,k} Pr(\Psi_{ji} = 0) \ \cdot \ Pr(\Phi_i = 0) \\ &= \Big(N_k(t)K_{ki}(\mathbf{N}(t))\Delta + \sigma(\Delta)\Big) \\ &\quad \cdot \prod_{j\in S, j\neq i,k} \Big(1 - N_j(t)K_{ji}(\mathbf{N}(t))\Delta + \sigma(\Delta)\Big) \\ &\quad \cdot \Big(1 + N_i(t)K_{ii}(\mathbf{N}(t))\Delta + \sigma(\Delta)\Big) \\ &= N_k(t)K_{ki}(\mathbf{N}(t))\Delta + \sigma(\Delta) \end{aligned}$$

In this case $\Omega = 1$. All cases when $k \in S, k \neq i$ result in $\Omega = 1$ with this tangible probability, and similar to the $\Omega = 0$ case, there are several other cases when $\Omega = 1$ but the associated probability is is $\sigma(\Delta)$.

Finally, when all Ψ_{ji} $(j \in S, j \neq i)$ equal to 0 and Φ_i equal to 1 we have $\Omega = -1$ with probability

$$\begin{aligned}\prod_{j\in S, j\neq i} Pr(\Psi_{ji} = 0) \ \cdot \ Pr(\Phi_i = 1) \\ = -N_i(t)K_{ii}(\mathbf{N}(t))\Delta + \sigma(\Delta) \ .\end{aligned}$$

The probability of other cases resulting in $\Omega = -1$ is $\sigma(\Delta)$ as well.

Putting together these cases we obtain the statement of the lemma. □

B Proof of Lemma 3

Proof. A binomially distributed random variable, Υ, with parameters n and p is defined as the result of n independent Bernoulli trials each of which results in 1 with probability p and 0 with probability $1 - p$. According to the strong law of large numbers

$$Pr\left(\lim_{n\to\infty} \frac{\Upsilon}{n} = E\left(\frac{\Upsilon}{n}\right)\right) = 1, \tag{28}$$

which means that Υ/n becomes deterministic (with probability 1) and it takes the value of its mean, which is independent of n.

Instead of (28), in a sloppy way, we simply write

$$\lim_{n\to\infty} \frac{\Upsilon}{n} = E\left(\frac{\Upsilon}{n}\right).$$

The same convergence rule applies for the weighted sum of independent binomially distributed random variables, like the one at the lhs of (10),

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N}\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right) \\ = E\left(\frac{1}{N}\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right)\right).\end{aligned}$$

Taking the expected value of the rhs we have

$$\begin{aligned}&\frac{1}{N}\ E\left(\sum_{j\in S, j\neq i} \Psi_{ji} - \Phi_i\right) \\ &= \sum_{j\in S, j\neq i} n_j(t)k_{ji}(\mathbf{n}(t))\Delta + n_i(t)k_{ii}(\mathbf{n}(t))\Delta \\ &= \sum_{j\in S} n_j(t)k_{ji}(\mathbf{n}(t))\Delta \ ,\end{aligned} \tag{29}$$

which results in the lemma. □

References

[1] F. Ball, R.K. Milne, I.D. Tame, and G.F. Yeo. Superposition of interacting aggregated continuous-time Markov chains. *Advances in Applied Probability*, 29:56–91, 1997.

[2] G. Bolch, S. Greiner, H. de Meer, and K.S. Trivedi. *Queueing Networks and Markov Chains*. Wiley, II Edition, 2006.

[3] J.Y. Le Boudec, D. McDonald, and J. Mundinger. A generic mean field convergence result for systems of interacting objects. In *4th Int Conf on Quantitative Evaluation of Systems - QEST2007*, 2007.

[4] P. Buchholz. Hierarchical Markovian models -symmetries and aggregation. *Performance Evaluation*, 22:93–110, 1995.

[5] P. Buchholz. Hierarchical structuring of superposed GSPNs. *IEEE Transactions Software Engineering*, 25:166–181, 1999.

[6] P. Buchholz and T. Dayar. Comparison of multilevel methods for Kronecker based Markovian representations. *Computing*, 73:349–371, 2004.

[7] M. Gribaudo, C.-F. Chiasserini, R. Gaeta, M. Garetto, D. Manini, and M. Sereno. A spatial fluid-based framework to analyze large-scale wireless sensor networks. In *IEEE International Conference on Dependable Systems and Networks, DSN2002*, 2005.

[8] M. Gribaudo, M. Telek, and A. Bobbio. Mean field methods in performance analysis. Technical report, Dip Informatica - Università Piemonte Orientale; www.di.unipmn.it/Tecnical-R/index.htm, 2008.

[9] J.M. Kelif and E. Altman. Downlink fluid model of CDMA networks. In *IEEE 61th Vehicular Technology Conference (VTC 2005)*, 2005.

[10] P. Kemper. Transient analysis of superposed GSPNs. *IEEE Trans Soft Engineering*, 25:182–193, 1999.

[11] M. Opper and D. Saad. *Advanced Mean Field Methods: Theory and Practice*. MIT University Press, 2001.

[12] B.D. Plateau and K. Atif. Stochastic automata network for modeling parallel systems. *IEEE Transactions on Software Engineering*, 17:1093–1108, 1991.

[13] B.D. Plateau and J.M. Fourneau. A methodology for solving Markov models of parallel systems. *Journal of Parallel and Distributed Computing*, 12:370–387, 1991.

[14] D. Zhang, D. Gatica-Perez, S. Bengio, and D. Roy. Learning influence among interacting Markov chains. In *Adv Neural Information Processing Systems (NIPS)*, 2005.