Simple Bounds on the Mean Cycle Time
in Acyclic Queueing Networks*

N. K. Krivulin
St.Petersburg State University

**Abstract**

Simple lower and upper bounds on mean cycle time in stochastic acyclic fork-join networks are derived using the $(\max,+)$-algebra approach.

# 1 Introduction

One of the problems of interest in the analysis of stochastic queueing networks is to evaluate the mean cycle time of a network. Both the mean cycle time and its inverse which can be regarded as a throughput, present performance measures commonly used to describe efficiency of the network operation.

It is frequently rather difficult to evaluate the mean cycle time exactly, even though the network under study is quite simple. To get information about the performance measure in this case, one can apply computer simulation to produce reasonable estimates. Another approach is to derive bounds on the mean cycle time (see [2, 4] for related examples).

The paper is concerned with the derivation of a lower and upper bounds on the mean cycle time for stochastic acyclic fork-join queueing networks [1, 4]. A useful way to represent dynamics of the networks is based on the $(\max,+)$-algebra approach [5, 4, 9]. We apply the $(\max,+)$-algebra dynamic representation proposed in [8] to get algebraic bounds and then exploit them to derive bounds in the stochastic case.

# 2 Algebraic Definitions and Results

The $(\max,+)$-algebra is an idempotent commutative semiring which is defined as $\mathbb{R}_{\max} = (\underline{\mathbb{R}},\oplus,\otimes)$ with $\underline{\mathbb{R}} = \mathbb{R}\cup\{\varepsilon\}$, $\varepsilon=-\infty$, and binary operations $\oplus$ and $\otimes$ defined as

$$x\oplus y=\max(x,y),\quad x\otimes y=x+y\quad \forall x,y\in\underline{\mathbb{R}}.$$

As it is easy to see, the operations $\oplus$ and $\otimes$ retain most of the properties of the ordinary addition and multiplication, including associativity, commutativity, and distributivity of multiplication over addition. However, the operation $\oplus$ is idempotent; that is, for any $x\in\underline{\mathbb{R}}$, one has $x\oplus x=x$.

There are the null and identity elements, namely $\varepsilon$ and $0$, to satisfy the conditions $x\oplus\varepsilon=\varepsilon\oplus x=x$, and $x\otimes 0=0\otimes x=x$, for any $x\in\underline{\mathbb{R}}$. The null element $\varepsilon$ and the operation $\otimes$ are related by the usual absorption rule involving $x\otimes\varepsilon=\varepsilon\otimes x=\varepsilon$.

*This work was supported in part by Russia Universities grant 4233

Non-negative power of any $x \in \mathbb{R}$ is defined as $x^{\otimes 0} = 0$, and $x^{\otimes q} = x \otimes \cdots \otimes x$, for any $q \geq 1$. Clearly, the new power $x^{\otimes q}$ corresponds to $qx$ in conventional algebra. In this paper, we will use the power notations only in the $(\max, +)$-algebra sense.

The $(\max, +)$-algebra of matrices is readily introduced in the regular way. Specifically, for any $(n \times n)$-matrices $X = (x_{ij})$ and $Y = (y_{ij})$, the entries of $U = X \oplus Y$ and $V = X \otimes Y$ are calculated respectively as

$$u_{ij} = x_{ij} \oplus y_{ij}, \quad \text{and} \quad v_{ij} = \sum_{k=1}^{n}{}_{\oplus}\, x_{ik} \otimes y_{kj},$$

where $\sum_{\oplus}$ stands for the iterated operation $\oplus$. As the null element, the matrix $\mathcal{E}$ with all entries equal to $\varepsilon$ is taken in the algebra, whereas the diagonal matrix $E = \mathrm{diag}(0, \ldots, 0)$ with the off-diagonal entries set to $\varepsilon$ presents the identity.

For any square matrix $X \neq \mathcal{E}$, one can define $X^{\otimes 0} = E$, and $X^{\otimes q} = X \otimes \cdots \otimes X$, for any $q \geq 1$. The norm of any matrix $X$ is defined as

$$\|X\|_{\oplus} = \sum_{i,j}{}_{\oplus}\, x_{ij} = \max_{i,j} x_{ij}.$$

Consider an $(n \times n)$-matrix $X$ with its entries $x_{ij} \in \underline{\mathbb{R}}$. It can be treated as an adjacency matrix of an oriented graph with $n$ nodes, provided each entry $x_{ij} \neq \varepsilon$ implies the existence of the arc $(i, j)$ in the graph, while $x_{ij} = \varepsilon$ does the lack of the arc.

It is easy to verify that for any integer $q \geq 1$, the matrix $X^{\otimes q}$ has its the entry $x_{ij}^{(q)} \neq \varepsilon$ if and only if there exists a path from node $i$ to node $j$ in the graph, which consists of $q$ arcs. Furthermore, if the graph associated with the matrix $X$ is acyclic, we have $X^{\otimes q} = \mathcal{E}$ for all $q > p$, where $p$ is the length of the longest path in the graph. Otherwise, provided that the graph is not acyclic, one can construct a path of any length, lying along circuits, and then it holds that $X^{\otimes q} \neq \mathcal{E}$ for all $q \geq 0$.

It is not difficult to verify the next statement.

**Proposition 1** *Let matrices $X_1, \ldots, X_k$ have a common associated acyclic graph, $p$ be the length of the longest path in the graph, and $X = X_1^{\otimes m_1} \otimes \cdots \otimes X_k^{\otimes m_k}$, where $m_1, \ldots, m_k$ are nonnegative integers.*

*If it holds that $m_1 + \cdots + m_k > p$, then $X = \mathcal{E}$.*

With Proposition 1, one can prove the next lemma.

**Lemma 1** *Let matrices $X_1, \ldots, X_k$ have a common associated acyclic graph, and $p$ be the length of the longest path in the graph.*

*If $\|X_i\|_{\oplus} \geq 0$ for all $i = 1, \ldots, k$, then for any nonnegative integers $m_1, \ldots, m_k$, it holds*

$$\left\| \prod_{i=1}^{k}{}_{\otimes} (E \oplus X_i)^{\otimes m_i} \right\|_{\oplus} \leq \left( \sum_{i=1}^{k}{}_{\oplus} \|X_i\|_{\oplus} \right)^{\otimes p},$$

*where $\prod_{\otimes}$ denotes the iterated operation $\otimes$.*

## 3 Elements of Probability

In this section we present some probabilistic results associated with $(\max, +)$-algebra concepts introduced above. We start with a lemma which states properties of the expectation with respect to the operations $\oplus$ and $\otimes$.

**Lemma 2** *Let* $\xi_1, \ldots, \xi_k$ *be random variables taking their values in* $\underline{\mathbb{R}}$*, and such that their expected values* $\mathbb{E}[\xi_i]$*,* $i = 1, \ldots, k$*, exist. Then it holds*

1. $\mathbb{E}[\xi_1 \oplus \cdots \oplus \xi_k] \geq \mathbb{E}[\xi_1] \oplus \cdots \oplus \mathbb{E}[\xi_k]$,
2. $\mathbb{E}[\xi_1 \otimes \cdots \otimes \xi_k] = \mathbb{E}[\xi_1] \otimes \cdots \otimes \mathbb{E}[\xi_k]$.

The next result [6] provides an upper bound for the expected value of the maximum of independent and identically distributed (i.i.d.) random variables.

**Lemma 3** *Let* $\xi_1, \ldots, \xi_k$ *be i.i.d. continuous random variables, and there exist* $\mathbb{E}[\xi_1] < \infty$ *and* $\mathbb{D}[\xi_1] < \infty$*. Then it holds*

$$\mathbb{E}\left[\sum_{i=1}^{k}{}_{\oplus}\, \xi_i\right] \leq \mathbb{E}[\xi_1] + \frac{k-1}{\sqrt{2k-1}}\sqrt{\mathbb{D}[\xi_1]}.$$

Consider a random matrix $X$ with its entries $x_{ij}$ taking values in $\underline{\mathbb{R}}$. We denote by $\mathbb{E}[X]$ the matrix obtained from $X$ by replacing each entry $x_{ij}$ by its expected value $\mathbb{E}[x_{ij}]$. With Lemma 2, it is easy to verify the next statement.

**Lemma 4** *For any random matrix* $X$*, it holds*

$$\mathbb{E}\left\|X\right\|_{\oplus} \geq \left\|\mathbb{E}[X]\right\|_{\oplus}.$$

# 4 Acyclic Fork-Join Networks

We consider a network with $n$ nodes and customers of a single class. The network topology is described by an oriented acyclic graph $\mathcal{G} = (\mathbf{N}, \mathbf{A})$, where $\mathbf{N} = \{1, \ldots, n\}$ represents the set of nodes, and $\mathbf{A} = \{(i,j)\} \subset \mathbf{N} \times \mathbf{N}$ is the set of arcs which determine the transition routes of customers. It is assumed that there are nodes in the graph which have no either incoming or outgoing arcs. Each node with no predecessors is assumed to represent an infinite external arrival stream of customers; provided that a node has no successors, it is considered as an output node intended to release customers from the network.

Each node of the network includes a server and infinite buffer which together present a single-server queue operating under the first-come, first-served queueing discipline. At the initial time, the server at each node is assumed to be free of customers, the buffers in nodes with no predecessors have infinite number of customers, whereas the buffers in the other nodes may have finite numbers of customers.

In addition to the usual service procedure, special join and fork operations [1] may be performed in a node respectively before and after service of a customer. The join operation is actually thought to cause each customer which comes into a node not to enter the buffer but to wait until at least one customer from every preceding node arrives. As soon as these customers arrive, they, taken one from each preceding node, are united to be treated as being one customer which then enters the buffer to become a new member of the queue.

The fork operation at a node is initiated every time the service of a customer is completed; it consists in giving rise to several new customers one for each succeeding nodes. These customers simultaneously depart the node, each being passed to separate node related to the first one. We assume that the execution of fork-join operations as well as the transition of customers within and between nodes require no time.

For the queue at node $i$, we denote the $k$th departure epoch by $x_i(k)$, and the service time of the $k$th customer by $\tau_{ik}$. We assume that $\tau_{ik} \geq 0$ are given parameters for all $i = 1, \ldots, n$, and $k = 1, 2, \ldots$, while $x_i(k)$ are considered as unknown state variables. With the condition that the network starts operating at time zero, it is convenient to set $x_i(0) = 0$ and $x_i(k) = \varepsilon$ for all $k < 0$, $i = 1, \ldots, n$. Finally, we denote the number of customers in the buffer at node $i$ at the initial time by $r_i$ with $0 \leq r_i \leq \infty$, and introduce $M = \max\{r_i |\ r_i < \infty;\ i = 1, \ldots, n\}$.

It has been shown in [8] that the dynamics of the network can be described by the state equation

$$\boldsymbol{x}(k) = \sum_{m=1}^{m}{}_{\oplus} A_m(k) \otimes \boldsymbol{x}(k-m), \tag{1}$$

with the state transition matrices

$$\begin{aligned} A_1(k) &= (E \oplus \mathcal{T}_k \otimes G_0^T)^{\otimes p} \otimes \mathcal{T}_k \otimes (E \oplus G_1^T), \\ A_m(k) &= (E \oplus \mathcal{T}_k \otimes G_0^T)^{\otimes p} \otimes \mathcal{T}_k \otimes G_m^T, \quad m = 2, \ldots, M, \end{aligned}$$

where $\mathcal{T}_k = \mathrm{diag}(\tau_{1k}, \ldots, \tau_{nk})$ with $\varepsilon$ as the off-diagonal entries, $G_m = (g_{ij}^m)$, $m = 0, 1, \ldots, M$, are matrices with their entries defined by the condition

$$g_{ij}^m = \begin{cases} 0, & \text{if } (i,j) \in \mathbf{A} \text{ and } m = r_j, \\ \varepsilon, & \text{otherwise,} \end{cases}$$

and $p$ is the length of the longest path in the graph associated with the matrix $G_0$.

## 5 Monotonicity Properties

In this section, a useful property of monotonicity is established which relates the system state vector $\boldsymbol{x}(k)$ to the initial numbers of customers $r_i$. It is actually shown that the entries of $\boldsymbol{x}(k)$ for all $k = 1, 2, \ldots$, do not decrease when the numbers $r_i$ with $0 < r_i < \infty$, $i = 1, \ldots, n$, are reduced to $0$.

We prove this assertion in two steps: first we consider the effect of reducing the numbers $r_i > 1$ to $1$, and then replace $r_i = 1$ with $r_i = 0$. Note that the change in the initial numbers of customers results only in modifications to adjacency matrices $G_m$. Specifically, reducing these numbers to $1$ leads us to new matrices $\widetilde{G}_1 = G_1 \oplus \cdots \oplus G_M$, and $\widetilde{G}_m = \mathcal{E}$ for all $m = 2, \ldots, M$, whereas $G_0$ remains unchanged.

**Lemma 5** *Let $\boldsymbol{x}(k)$ be determined by (1). Suppose that the vector $\widetilde{\boldsymbol{x}}(k)$ satisfies the dynamic equation*

$$\widetilde{\boldsymbol{x}}(k) = \widetilde{A}_1(k) \otimes \widetilde{\boldsymbol{x}}(k-1), \quad \widetilde{\boldsymbol{x}}(0) = 0,$$

*with $\widetilde{A}_1(k) = (E \oplus \mathcal{T}_k \otimes G_0^T)^{\otimes p} \otimes \mathcal{T}_k \otimes (E \oplus \widetilde{G}_1^T)$, where $\widetilde{G}_1 = G_1 \oplus \cdots \oplus G_M$.*

*Then for all $k = 1, 2, \ldots$, it holds*

$$\boldsymbol{x}(k) \leq \widetilde{\boldsymbol{x}}(k).$$

The next lemma shows that in the system with finite initial numbers $r_i$ equal either to $0$ or $1$, the entries of its transition matrix $A_1$ do not decrease as all these numbers are set to $0$. Clearly, this also involves nondecrease in the system state vector $\boldsymbol{x}(k)$.

**Lemma 6** *For all* $k = 1, 2, \ldots$, *it holds*

$$A_1(k) \leq \widetilde{A}(k)$$

*with* $\widetilde{A}(k) = (E \oplus \mathcal{T}_k \otimes \widetilde{G}^T)^{\otimes q} \otimes \mathcal{T}_k$, *where* $\widetilde{G} = G_0 \oplus G_1$, *and* $q$ *is the length of the longest path in the graph associated with the matrix* $\widetilde{G}$.

Lemmas 5 and 6 can be summarized as follows.

**Theorem 1** *In the acyclic fork-join queueing network, reducing the initial numbers of customers from any fixed values to* 0 *does not decrease the entries of the system state vector* $\boldsymbol{x}(k)$ *for all* $k = 1, 2, \ldots$

# 6 Service Cycle Time

We consider the evolution of the system (1) as a sequence of service cycles: the 1st cycle starts at the initial time, and it is terminated as soon as all the servers in the network complete their 1st service, the 2nd cycle is terminated as soon as the servers complete their 2nd service, and so on. The completion time of the $k$th cycle can be represented as

$$\max_i x_i(k) = \|\boldsymbol{x}(k)\|_\oplus .$$

The next statement based on Lemma 1 and Theorem 1 provides simple algebraic lower and upper bounds on the $k$th cycle completion time.

**Lemma 7** *For all* $k = 1, 2, \ldots$, *it holds*

$$\left\| \sum_{i=1}^{k} \mathcal{T}_i \right\|_\oplus \leq \|\boldsymbol{x}(k)\|_\oplus \leq \sum_{i=1}^{k} \|\mathcal{T}_i\|_\oplus + p \left( \sum_{i=1}^{k}{}_\oplus \|\mathcal{T}_i\|_\oplus \right).$$

In many applications, one is normally interested in investigating the steady-state mean cycle time; that is, the limit of $\|\boldsymbol{x}(k)\|_\oplus / k$ as $k$ tends to $\infty$. We will consider this problem with relation to the stochastic network models in the next section.

# 7 Stochastic Networks

Suppose that for each node $i = 1, \ldots, n$, the service times $\tau_{i1}, \tau_{i2}, \ldots$, form a sequence of i.i.d. continuous non-negative random variables with $\mathbb{E}[\tau_{ik}] < \infty$ and $\mathbb{D}[\tau_{ik}] < \infty$ for all $k = 1, 2, \ldots$. With these conditions, $\mathcal{T}_k$ are i.i.d. random matrices, whereas $\|\mathcal{T}_k\|_\oplus$ present i.i.d. random variables with $\mathbb{E}\,\|\mathcal{T}_k\|_\oplus < \infty$ and $\mathbb{D}\,\|\mathcal{T}_k\|_\oplus < \infty$ for all $k = 1, 2, \ldots$.

In the analysis of the mean cycle time of the system, one first has to convince himself that the limit

$$\lim_{k \to \infty} \frac{1}{k} \|\boldsymbol{x}(k)\|_\oplus = \gamma \tag{2}$$

exists. A standard technique to verify the existence of the above limit is based on the Subadditive Ergodic Theorem proposed in [7]. One can find examples of the implementation of the theorem in the $(\max, +)$-algebra framework in [3, 4].

With the above probabilistic conditions, it is not difficult to apply the theorem in the same way as in [3, 4] to prove existence of the limit in (2) for the system (1). In more exact terms, the theorem allows one to prove that this limit exists with the probability one, and, at the same time, it holds

$$\lim_{k\to\infty} \frac{1}{k} \mathbb{E} \, \|\boldsymbol{x}(k)\|_{\oplus} = \gamma. \qquad (3)$$

Now we present our main result which offers simple bounds on the mean cycle time $\gamma$.

**Theorem 2** *In the stochastic dynamical system (1) the mean cycle time* $\gamma$ *satisfies the double inequality*

$$\|\mathbb{E}[\mathcal{T}_1]\|_{\oplus} \leq \gamma \leq \mathbb{E} \, \|\mathcal{T}_1\|_{\oplus} \, . \qquad (4)$$

**Proof:** Note that according to the Subadditive Ergodic Theorem, we may represent the mean cycle time $\gamma$ as (3).

Let us first prove the left inequality in (4). From Lemmas 7 and 4, we have

$$\frac{1}{k} \mathbb{E} \, \|\boldsymbol{x}(k)\|_{\oplus} \geq \frac{1}{k} \mathbb{E} \left\| \sum_{i=1}^{k} \mathcal{T}_i \right\|_{\oplus} \geq \left\| \frac{1}{k} \sum_{i=1}^{k} \mathbb{E}[\mathcal{T}_i] \right\|_{\oplus} = \|\mathbb{E}[\mathcal{T}_1]\|_{\oplus} \, ,$$

independently of $k$.

With the upper bound offered by Lemma 7, we get

$$\frac{1}{k} \mathbb{E} \, \|\boldsymbol{x}(k)\|_{\oplus} \leq \mathbb{E} \, \|\mathcal{T}_1\|_{\oplus} + \frac{p}{k} \mathbb{E} \left[ \sum_{i=1}^{k}{}_{\oplus} \, \|\mathcal{T}_i\|_{\oplus} \right] .$$

From Lemma 3, the second term on the right-hand side may be replaced by that of the form

$$\frac{p}{k} \left( \mathbb{E} \, \|\mathcal{T}_1\|_{\oplus} + \frac{k-1}{\sqrt{2k-1}} \sqrt{\mathbb{D} \, \|\mathcal{T}_1\|_{\oplus}} \right) ,$$

which tends to 0 as $k \to \infty$. □

# References


1. *Baccelli F., Makowski A.M.* Queueing Models for Systems with Synchronization Constraints // Proc. IEEE. 1989. V. 77. N 1. P. 138-160.

2. *Baccelli F., Konstantopoulos P.* Estimates of Cycle Times in Stochastic Petri Nets // Applied Stochastic Analysis: Lecture Notes in Control and Information Sciences. V. 177. New York, 1991. P. 1-20.

3. *Baccelli F.* Ergodic Theory of Stochastic Petri Networks // Ann. Probab. 1992. V. 20. N 1. P. 375-396.

4. *Baccelli F., Cohen G., Olsder G.J., Quadrat J.-P.* Synchronization and Linearity. Chichester, 1992.

5. *Cuninghame-Green R.A.* Minimax Algebra: Lecture Notes in Economics and Mathematical Systems. V. 166. Berlin, 1979.

6. *Hartly H.O., David H.A.* Universal Bounds for Mean Range and Extreme Observations // Ann. Math. Statist. 1954. V. 25. P. 85-99.

7. *Kingman J.F.C.* Subadditive Ergodic Theory // Ann. Probab. 1973. V. 1. P. 883-909.

8. *Krivulin N.K.* Max-Plus Algebra Models of Queueing Networks // Proc. International Workshop on Discrete Event Systems (WODES'96), University of Edinburgh. London, 1996. P. 76-81.

9. *Maslov V.P., Kolokoltsov V.N.* Idempotent Analysis and Its Applications to Optimal Control Theory. Moscow, 1994.