

Article

Not peer-reviewed version

Identifying Multiple Randomness in Random Experiments: Definition and Examples

[Guang-Liang Li](#) *

Posted Date: 4 August 2025

doi: 10.20944/preprints202507.2444.v2

Keywords: random phenomena with multiple randomness; stochastic process; counting process; queuing theory



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Identifying Multiple Randomness in Random Experiments: Definition and Examples

Guang-Liang Li 

University of Hong Kong, Hong Kong, China; glli@eee.hku.hk

Abstract

Multiple randomness is a random phenomenon appeared in experiments concerning practical applications of sciences and engineering disciplines. But phenomena with multiple randomness are ignored in the existing literature, because correctly derived results in mathematics are used to interpret, incorrectly, outcomes obtained by experiments designed to solve practical problems in the real world, leading to questionable theorems in practice. Multiple randomness is characterized by unique properties of the sample space consisting of all possible outcomes obtained by the corresponding random experiment and differs essentially from any known phenomenon observed. By introducing general results concerning multiple randomness in a random experiment performed to count the number of “events” and providing specific examples to illustrate the properties of multiple randomness in counting processes, this study aims to demonstrate the existence of such phenomena. The specific examples are popular models taken from queuing theory, which may help the reader to understand the general results. It is reasonable to expect more phenomena with multiple randomness to be identified in other stochastic processes and in other application fields of sciences and engineering disciplines, which may help scientists and engineers to explain weird phenomena and solve puzzling problems in the real world.

Keywords: random phenomena with multiple randomness; stochastic process; counting process; queuing theory

MSC: 60G20, 90B15, 90B22, 60K25

1. Introduction

In experiments concerning practical applications of sciences and engineering disciplines, phenomena with multiple randomness appear naturally but are ignored in the existing literature, because correctly derived results in mathematics are used to interpret, incorrectly, outcomes obtained by experiments designed to solve practical problems in the real world, leading to questionable theorems in practice. By introducing multiple randomness identified in an experiment performed to count the number of “events”, this study aims to demonstrate the existence of such phenomena.

Many scientists and engineers use counting processes to study interesting phenomena in the real world. Let $N(t)$ be a counting process. As a stochastic process, $N(t)$ represents the total number of events occurred up to time $t \geq 0$. If the j th event occurs at instant τ_j , where $j \geq 1$, the inter-event time $\tau_{j+1} - \tau_j$ is a random variable. If an experiment is performed to count the number of events, the outcome cannot be determined in advance; counting the number of events is a random experiment; its sample space Ω is the set of all possible outcomes, which may take different forms. In this study, Ω consists of all possible sequences of instants at which the events occur, i.e.,

$$\Omega = \{(\tau_n)_{n \geq 1}, \tau_{n+1} > \tau_n\}.$$

If $\tau_{j+1} - \tau_j$ for every j is defined on the *whole* sample space, i.e., a value can be assigned to $\tau_{j+1} - \tau_j$ at each $\omega \in \Omega$, then $N(t)$ is a *usual counting process*. For usual counting processes, such as a renewal process, all inter-event times can be characterized by a marginal distribution.

There are also random phenomena modeled by counting processes different from any known phenomenon. The difference is this: for any j , the inter-event time $\tau_{j+1} - \tau_j$ is defined only on some *proper subset* of Ω and cannot be characterized by a marginal distribution. Such counting processes differ essentially from any known stochastic process in the existing literature. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probability space of the random experiment performed to count the number of events, where $\mathcal{A}$ represents the σ -algebra of subsets of Ω , and the probability measure $\mathbb{P}$ is defined on the measurable space $(\Omega, \mathcal{A})$. Denote by $\mathbb{N}$ the set of all positive integers. Let $m \geq 1$ be the total number of distributions for inter-event times. Any inter-event time can only follow one and only one of such distributions. If $m = 1$, then for all $j \geq 1$, the inter-event times $\tau_{j+1} - \tau_j$ can be characterized by a marginal distribution, and $N(t)$ degenerates into a usual counting process. For an arbitrarily *given* $j \in \mathbb{N}$, write

$$\Omega = \Omega_1(j) \cup \Omega_2(j) \cdots \cup \Omega_m(j)$$

where

$$\Omega_i(j) = \{\omega \in \Omega : (\tau_{j+1} - \tau_j)(\omega) = T_i(\omega)\}, \mathbb{P}[\Omega_i(j)] > 0, \sum_{i=1}^m \mathbb{P}[\Omega_i(j)] = 1.$$

Denote by F_i the distribution of T_i in the expression of $\Omega_i(j)$. If $k' \neq k$, where $1 < k' \leq m$ and $1 < k \leq m$, then $F_{k'} \neq F_k$, thus $\Omega_{k'}(j) \cap \Omega_k(j) = \emptyset$, i.e., $\{\Omega_i(j), i = 1, 2, \dots, m\}$ is a partition of Ω . Let $\omega \in \Omega$ be an arbitrarily *given* sample point. Write

$$\mathbb{N} = M_1(\omega) \cup M_2(\omega) \cdots \cup M_m(\omega)$$

where

$$M_i(\omega) = \{j \in \mathbb{N} : (\tau_{j+1} - \tau_j)(\omega) = T_i(\omega)\}, i \in \{1, 2, \dots, m\}$$

constitute a *random partition* of $\mathbb{N}$. The partition is random, because $M_i(\omega)$ depends on the sample point. At the given ω , if $k' \neq k$, then $M_{k'}(\omega) \cap M_k(\omega) = \emptyset$.

Definition 1. If Ω possesses the properties given above, $N(t) \geq 0$ is said to be a counting process with multiple (or m -fold) randomness for $m > 1$. The corresponding inter-event times constitute a stochastic sequence $(\tau_{j+1} - \tau_j)_{j \geq 1}$ with m -fold randomness.

According to the above definition, for a counting process with m -fold randomness, the inter-event time $\tau_{j+1} - \tau_j$ is not defined on the whole sample space for any j and cannot be characterized by a marginal distribution. The distribution of $\tau_{j+1} - \tau_j$ is not determined by a joint distribution of random variables on Ω ; it is determined uniquely by the random experiment in question. As terms of the stochastic sequence $(\tau_{j+1} - \tau_j)_{j \geq 1}$, the inter-event times are not independent random variables. Furthermore, immediately from Definition 1, $(\tau_{j+1} - \tau_j)_{j \geq 1}$ is not a stationary sequence.

Counting processes with multiple randomness appear naturally in practical applications, such as queuing theory applied to different fields (for example, see [1–3]). Unfortunately, they have not been identified and are all mistaken for their usual counterparts. The best way of demonstrating their existence and illustrating their properties is to provide examples. To this end, several well-known queuing models will be considered. By identifying multiple randomness in the queuing models, an inconsistency in queuing theory concerning departure processes of stable queues in steady state and product-form equilibrium distributions of queuing networks can be readily resolved.

A lemma and its corollary concerning the departure process from a stable, general single-server queue in steady state are proved: the instants of departures from this queue constitute a counting process with two-fold randomness (i.e., $m = 2$), or equivalently, the corresponding inter-departure times form a stochastic sequence with two-fold randomness (Section 2). Based on the lemma and its

corollary, the departure process of a stable M/M/1 queue in steady state is revisited (Section 3), and stability of Jackson networks of queues is scrutinized (Section 4). The article is concluded in Section 5.

2. Two-Fold Randomness in Counting Processes: General Results

Consider departures from a stable, general single-server queue (i.e., a GI/GI/1 system) in steady state. This queuing model has an infinite waiting room and a work-conserving server with a finite service capacity defined by the maximum rate at which the server can perform work. The meaning of “work-conserving” is this: the server will not stand idle when there is unfinished work in the system. Customers arrive at this system according to a renewal process and are served one at a time. Times between successive arrivals are independent and identically distributed (i.i.d.) random variables with a finite mathematical expectation, and so are service times of customers. Inter-arrival times and service times are also mutually independent. For such a queue, a customer leaves the system if and only if the customer has been served. The mean inter-arrival time is greater than the mean service time. Hence the queue is stable and can reach its steady state as time approaches infinity. For the purpose of this study, it is not necessary to assume a specific queuing discipline.

According to the literature, times between consecutive departures from a stable queue in steady state follow a marginal distribution [4]. However, the existence of such a marginal distribution is only an unjustified assumption taken for granted without verification. As can be readily seen in this section, for systems modeled by a stable GI/GI/1 queue, inter-departure times between customers consecutively leaving from the queue cannot be characterized by a marginal distribution, even if the queue is in steady state. Usually, properties of the arrival process and stability of the GI/GI/1 queue are established by using the strong law of large numbers. Results of this kind are true for every $\omega \in \Omega$, with the exception at most of a set $\mathcal{N} \in \mathcal{A}$ where $\mathbb{P}(\mathcal{N}) = 0$. Such results hold with probability one or almost surely. Removing sample points in $\mathcal{N}$ will not change anything when the queue is already in steady state, and Ω can be regarded as the sample space of an ideal random experiment performed after the queue has reached its steady state.

Let c_j represent the j th customer served. Using the notation introduced in Section 1, the event “ c_j departs from the queue” occurs at τ_j . Denote by Q_j the queue size immediately after the j th departure. By definition, “queue size” refers to the total number of customers in the queue (including the customer in service). Let T_1 and T_2 be the times between the j th and the $(j+1)$ th departures when $Q_j = 0$ and $Q_j > 0$, respectively.

$$T_1 = I_j + S_{j+1} \quad (2.1)$$

$$T_2 = S_{j+1} \quad (2.2)$$

where I_j is the idle time spent by the server waiting for the arrival of a customer, and S_{j+1} is the service time of c_{j+1} . All arrived customers need to be served; their service times are not zero.

By definition, a random variable is a measurable real-valued function; its domain can be the whole sample space or a proper subset of the sample space. To define a random variable U on the whole sample space, a value must be assigned to U at each $\omega \in \Omega$; To define a random variable on a proper subset of Ω , a value must be assigned to this random variable at each sample point in the subset. Denote by $\mathcal{B}$ the Borel algebra of subsets in the real line. For a random variable on the whole sample space, such as U , its (marginal) distribution is defined by

$$P_U(B) = \mathbb{P}(U \in B)$$

where $B \in \mathcal{B}$ is arbitrary. When it is necessary to emphasize the connection between U and $\mathbb{P}$ on $(\Omega, \mathcal{A})$, the right-hand side of the above equation can be used to express the distribution of U directly. Similarly, components of a random vector or terms of a stochastic sequence are random variables on Ω . By definition, the components of a random vector (or the terms of a stochastic sequence) must take their values at the same $\omega \in \Omega$. If a random vector has been defined, then the joint distribution of its components is fixed, and the marginal distribution of each component is determined by the joint

distribution. Based on a given random vector, some new random variables may be constructed on proper subsets of Ω . For example, let (U, V) be a random vector. The joint distribution is $P_{U,V}$, which determines P_U and P_V , the marginal distributions of U and V . Based on this random vector, a random variable W may be constructed on $\Theta \subset \Omega$ with $\mathbb{P}(\Theta) > 0$, such that $W(\omega) = V(\omega)$ for $\omega \in \Theta$; some values taken by U specify Θ . Thus, W follows a conditional distribution P_W determined by $P_{U,V}$.

So far, two types of random variables have been mentioned; U and V are defined on the whole sample space, but W is defined on a proper subset of the sample space and follows a conditional distribution. There also exist random variables different from those mentioned above. Random variables of this type are not components of a random vector, and their distributions are not determined by a joint distribution. Times between successive departures from the GI/GI/1 queue are random variables of this type. Proper subsets of Ω are their domains. A *chronological order of events experienced by customers* determines their distributions and properties. This chronological order is not determined by properties of the queuing model; it is determined by physical systems modeled by the queue. For a work-conserving system with an infinite waiting room, events experienced by each customer occur naturally as follows:

- First, a customer arrives at the queue.
- Upon arrival,
 - the customer receives service immediately if the server is idle;
 - otherwise the customer has to wait in line.
- Finally, after being served, the customer departs.

Consequently, for an arbitrarily given j , if the server becomes idle immediately after c_j leaves, the time between the departures of c_j and c_{j+1} is the sum of an idle time and a service time, as shown by Eq.(2.1); otherwise the inter-departure time is a service time, see Eq.(2.2).

Lemma 1. *If a stable GI/GI/1 queue is in steady state, and if the server has a finite service capacity, then for an arbitrarily given $j \in \mathbb{N}$,*

- (a) $\tau_{j+1} - \tau_j$ cannot be expressed by any single, fixed random variable defined on Ω or characterized by a marginal (i.e., unconditional) distribution;
- (b) $(Q_j, \tau_{j+1} - \tau_j)$ is not a random vector, and hence $\tau_{j+1} - \tau_j$ has no distributions conditional on values taken by Q_j .

Corollary 1. *The inter-departure times constitute a stochastic sequence $(\tau_{j+1} - \tau_j)_{j \geq 1}$ with two-fold randomness.*

To understand Lemma 1 and its corollary, the reader may consider how to answer the following question: if $\tau_{j+1} - \tau_j$ could be expressed by a single, fixed random variable Z_j defined on Ω , what would be the value of Z_j at ω corresponding to $Q_j(\omega)$? At any $\omega \in \Omega$, the value of Q_j equals either zero or a positive integer. Whatever value Q_j takes at ω , it is impossible to assign any value to Z_j at ω corresponding to $Q_j(\omega)$. In contrast to $\tau_{j+1} - \tau_j$, service times of customers, the queue size Q_j , and times between consecutive arrivals are all random variables on Ω . There is a subtlety here, however. For example, when playing the role of an inter-departure time T_2 defined on $\Omega_2(j)$ (see below), S_{j+1} is no longer a random variable on Ω .

Proof. Because of the chronological order, the corresponding sample space Ω of the general queuing model possesses the properties required by a counting process with two-fold randomness. To see this, write

$$M_1(\omega) = \{j \in \mathbb{N} : Q_j(\omega) = 0\}$$

$$M_2(\omega) = \{j \in \mathbb{N} : Q_j(\omega) > 0\}$$

$$\Omega_1(j) = \{\omega \in \Omega : Q_j(\omega) = 0\}$$

and

$$\Omega_2(j) = \{\omega \in \Omega : Q_j(\omega) > 0\}.$$

Clearly, at an arbitrarily given $\omega \in \Omega$, $\{M_1(\omega), M_2(\omega)\}$ is a random partition of $\mathbb{N}$, and it is not difficult to verify $\Omega_1(j) \cap \Omega_2(j) = \emptyset$ for an arbitrarily given $j \geq 1$: if $\Omega_1(j) \cap \Omega_2(j) \neq \emptyset$, then for the given j , there would exist a sample point $\omega \in \Omega_1(j) \cap \Omega_2(j)$; when c_j departs, the server would be both idle and busy, which is impossible. Thus $\{\Omega_1(j), \Omega_2(j)\}$ is a partition of the sample space.

$$\Omega_1(j) \cup \Omega_2(j) = \Omega.$$

Because the queue is already in steady state, $\mathbb{P}[\Omega_1(j)] = 1 - \rho > 0$, $\mathbb{P}[\Omega_2(j)] = \rho > 0$, and

$$\mathbb{P}[\Omega_1(j)] + \mathbb{P}[\Omega_2(j)] = 1$$

where ρ is called the utilization factor. Consequently, at an arbitrarily given $\omega \in \Omega$, either $\omega \in \Omega_1(j)$ and

$$\left. \begin{aligned} Q_j(\omega) &= 0 \\ \tau_{j+1}(\omega) - \tau_j(\omega) &= T_1(\omega) \end{aligned} \right\} \quad (2.3)$$

or $\omega \in \Omega_2(j)$ and

$$\left. \begin{aligned} Q_j(\omega) &> 0 \\ \tau_{j+1}(\omega) - \tau_j(\omega) &= T_2(\omega). \end{aligned} \right\} \quad (2.4)$$

As shown by Eq.(2.1) and Eq.(2.2), T_1 and T_2 are well-defined random variables; their distributions, P_{T_1} and P_{T_2} , are determined by the chronological order of events experienced by customers, as implied by Eq.(2.3) and Eq.(2.4). For an arbitrarily fixed $B \in \mathcal{B}$,

$$P_{T_1}(B) = \mathbb{P}[T_1 \in B | \Omega_1(j)], \quad P_{T_2}(B) = \mathbb{P}[T_2 \in B | \Omega_2(j)].$$

Either Eq.(2.3) or Eq.(2.4) must hold *exclusively* at the given sample point $\omega \in \Omega$, and for the given j , $\tau_{j+1} - \tau_j$ can only be described by T_1 or T_2 . To see this, suppose to the contrary: there is a random variable $Z_j = \tau_{j+1} - \tau_j$ on Ω , which can take values at the same sample point ω corresponding to $Q_j(\omega)$. Consequently,

$$\{\Omega_1(j), T_1 \in B\} = \{\Omega_1(j), Z_j \in B\} \quad (2.5)$$

$$\{\Omega_2(j), T_2 \in B\} = \{\Omega_2(j), Z_j \in B\}. \quad (2.6)$$

Because $\{\Omega_1(j)\}$ is independent of events concerning future departures after c_j leaves, such as $\{T_1 \in B\}$ and $\{Z_j \in B\}$, Eq.(2.5) implies

$$\mathbb{P}[\Omega_1(j)]P_{T_1}(B) = \mathbb{P}[\Omega_1(j)]P_{Z_j}(B)$$

where P_{Z_j} is the distribution of Z_j . Similarly, $\{\Omega_2(j)\}$ is independent of $\{T_2 \in B\}$ and $\{Z_j \in B\}$, so Eq.(2.6) implies

$$\mathbb{P}[\Omega_2(j)]P_{T_2}(B) = \mathbb{P}[\Omega_2(j)]P_{Z_j}(B).$$

Because $\mathbb{P}[\Omega_1(j)] = 1 - \rho > 0$ and $\mathbb{P}[\Omega_2(j)] = \rho > 0$ for all j , treating $\tau_{j+1} - \tau_j$ as Z_j on Ω leads to

$$P_{T_1}(B) = P_{T_2}(B) = P_{Z_j}(B).$$

This is absurd. The absurdity shows $\{Z_j \in B\} \notin \mathcal{A}$. Consequently, $\tau_{j+1} - \tau_j$ cannot be described by any single, fixed random variable on Ω or characterized by a marginal distribution.

By definition, each component of a random vector (or each term of a stochastic sequence) is a random variable on Ω , and all the components (or terms) must take values at the *same* $\omega \in \Omega$. Although Q_j is a random variable on Ω , $\tau_{j+1} - \tau_j$ and Q_j cannot form a random vector. Therefore, $\tau_{j+1} - \tau_j$ is not eligible to have distributions conditional on values taken by Q_j , and $\tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ cannot form a sequence of random variables on Ω ; they constitute a stochastic sequence with two-fold randomness. $\square$

The conditions required by Lemma 1 exclude two conceivable scenarios for $\tau_{j+1} - \tau_j$ to be a single, fixed random variable defined on Ω . One is the worst scenario, and the other is an ideal scenario. The worst scenario is a queue with $\mathbb{P}[\Omega_1(j)] = 0$ for all j , i.e., the queue is unstable. Because $\Omega_1(j) \cup \Omega_2(j) = \Omega$, and because $\mathbb{P}[\Omega_1(j)] = 0$,

$$\mathbb{P}[\Omega_1(j) \cup \Omega_2(j)] = \mathbb{P}[\Omega_2(j)] = 1.$$

Thus, with probability one, when each customer departs from the system, the queue is not empty, which implies $\tau_{j+1} - \tau_j = S_{j+1}$ for all j . In other words, the idle time I_j vanishes identically on Ω , making the server always busy almost surely. Consequently, the queue will never become empty, and the queue size will approach infinity with probability one.

The ideal scenario is a queue with $\mathbb{P}[\Omega_2(j)] = 0$ for all j , i.e., the queue is always empty almost surely, because $\mathbb{P}[\Omega_2(j)] = 0$ for all j implies

$$\mathbb{P}[\Omega_1(j) \cup \Omega_2(j)] = \mathbb{P}[\Omega_1(j)] = 1.$$

In other words, when each customer departs from the system, the server becomes idle almost surely, which implies $\tau_{j+1} - \tau_j = I_j$ for every j . That is, the server must have an infinite service capacity; this is the only way to make the service times of customers vanish identically. Consequently, I_j becomes an inter-arrival time with probability one.

The scenarios above illustrate two extreme situations. If a stable queue is in steady state, then $\mathbb{P}[\Omega_1(j)] > 0$ and $\mathbb{P}[\Omega_2(j)] > 0$ for each j . Consequently, $(\tau_{j+1} - \tau_j)_{j \geq 1}$ is a stochastic sequence with two-fold randomness and differs essentially from any sequence of random variables defined on the whole sample space Ω . To illustrate the difference, consider an arbitrarily *given* $\omega \in \Omega$. At this sample point,

- (i) if $j \in M_1(\omega)$ (i.e., immediately after c_j departs, there is no unfinished work in the system), then $\tau_{j+1} - \tau_j$ equals $T_1(\omega)$; otherwise the server is busy, and $\tau_{j+1} - \tau_j$ takes $T_2(\omega)$ as its value;
- (ii) thus, $(\tau_{j+1} - \tau_j)_{j \geq 1}$ can be divided into two subsequences randomly as follows.

$$[\tau_{j+1}(\omega) - \tau_j(\omega)]_{j \geq 1} = [T_1(\omega)]_{j \in M_1(\omega)} \cup [T_2(\omega)]_{j \in M_2(\omega)}.$$

Having a finite service capacity, the server is either busy or idle with a positive probability, thus realistic inter-departure times always fall into two categories. In steady state, the probability for the server to be busy or idle is fixed, and the values of inter-departure times are given by either $T_1(\omega)$ or $T_2(\omega)$ according to whether $\omega \in \Omega_1(j)$ (i.e., the server is idle) or $\omega \in \Omega_2(j)$ (i.e., the server is busy) immediately after c_j departs, see Eq.(2.1) and Eq.(2.2). In the literature, the distribution of T_1 or T_2 is interpreted as the distribution of $\tau_{j+1} - \tau_j$ conditional on values taken by Q_j . According to such interpretation, a marginal distribution of the inter-departure times could be constructed based on the distributions of T_1 and T_2 .

However, the above interpretation is questionable. By Lemma 1, for each $j \in \mathbb{N}$, $\tau_{j+1} - \tau_j$ cannot be expressed as a random variable defined on the whole sample space Ω or characterized by a marginal distribution; it is problematic to treat the distribution of T_1 or T_2 as the distribution of $\tau_{j+1} - \tau_j$ conditional on values taken by Q_j , because $(Q_j, \tau_{j+1} - \tau_j)$ is not a random vector; its components cannot take values at the *same* sample point ω . As shown in the proof of Lemma 1, treating $(Q_j, \tau_{j+1} - \tau_j)$ as a random vector leads to a contradiction. According to Corollary 1, the

terms of $(\tau_{j+1} - \tau_j)_{j \geq 1}$ are not random variables on Ω ; they are terms of a stochastic sequence with two-fold randomness. The general results presented in this section may be better understood if specific examples are provided. A stable M/M/1 queue in steady state may serve as such an example, which will be considered in the following section.

3. Two-Fold Randomness in Counting Processes: Specific Examples

A stable M/M/1 queue in steady state is the simplest instance of the GI/GI/1 queue. Customers arrive at this system according to a Poisson process; service times are mutually independent and follow a common exponential distribution. According to Burke's theorem [5], times between successive departures from this queue in steady state are mutually independent and follow a marginal distribution identical to the distribution of inter-arrival times. Let λ be the parameter of this distribution. Clearly, Burke's theorem contradicts Lemma 1 and Corollary 1. In the following, Burke's original proof given in [5] (Subsection 3.1), Reich's proof based on time reversibility given in [6] (Subsection 3.2), and experimental results obtained by simulation concerning the departure process from the M/M/1 queue (Subsection 3.3) will be scrutinized. As can be readily seen, although Burke's theorem is considered as a well-established result in the literature, an inevitable conclusion can be derived: Burke's theorem is questionable. More examples concerning queues in tandem and networks of queues will be considered in Section 4.

3.1. Burke's Proof Based on Differential Equations

Burke considered "the length of an arbitrary inter-departure interval" in his proof [5]. Expressed by the notation in the present paper, what Burke called "an arbitrary inter-departure interval" is (τ_j, τ_{j+1}) for an arbitrary j . Denote by $Q(t)$ the state of the queue, which represents the queue size. Burke's proof begins with calculating the probability of an event $\{Q(t) = k, \tau_{j+1} - \tau_j > t\}$, where t is an instant after τ_j , and τ_j is taken to be the instant of the last previous departure. By comparing $\tau_{j+1} - \tau_j$ with t , Burke chose τ_j as the origin on the time axis, namely, c_j departs at $t = 0$. During the interval (τ_j, τ_{j+1}) , the number of arrived customers is at least one. Denote by t_1 the first arrival instant in this interval, where $0 < t_1 < \tau_{j+1}$.

Burke grounded his proof on an unjustified assumption unconsciously: $\tau_{j+1} - \tau_j$ is a single, fixed random variable Z_j defined on the whole sample space and can take values together with $Q(t)$ at any $\omega \in \Omega$. Burke did not mention this assumption in his paper; he took the assumption for granted. However, this assumption is false as shown in Section 2. Under this false assumption, Burke treated $[Q(t), \tau_{j+1} - \tau_j]$ as a random vector. In Burke's proof, a set of differential equations governing the probability of $\{Q(t) = k, Z_j > t\}$ is established. Solving the equations yields [5]

$$\mathbb{P}[Q(t) = k, Z_j > t] = \mathbb{P}[(Q(t) = k)]\mathbb{P}(Z_j > t), \quad k = 0, 1, \dots$$

Burke gave the following interpretation of the solutions he obtained by solving the equations:

- $\{Q(t) = k\}$ and $\{Z_j > t\}$ are independent events for any $k \geq 0$;
- $\mathbb{P}[Q(t) = k]$ is the equilibrium probability of $\{Q(t) = k\}$, which actually does not depend on t ;
- $\mathbb{P}(Z_j > t)$ is the distribution of inter-departure times and identical to the distribution of inter-arrival times.

$$\mathbb{P}(Z_j > t) = e^{-\lambda t}.$$

However, although the differential equations established by Burke and their solutions are mathematically correct, the false assumption underlying Burke's proof makes his theorem questionable; the theorem is nothing but the interpretation of the mathematical results: the equations and their solutions. The mathematical results belong to pure mathematics; Burke's theorem belongs to practical applications of the mathematical results. Interpretations of correct results in pure mathematics may be incorrect in practical applications, which may lead to questionable theorems in practice.

To see this, consider $Q(t) = 0$ first, i.e., the server is idle at time $t > 0$. If $Q(t) = 0$, then $0 < t < t_1$, and

$$\{Q(t) = 0, Z_j > t\} \subset \{Q(0) = 0, Z_j > t\}. \quad (3.1)$$

In other words, the occurrence of $\{Q(0) = 0, Z_j > t\}$ implies the occurrence of $\{Q(t) = 0, Z_j > t\}$: if $\{Q(t) = 0, Z_j > t\}$ occurs, then $\{Q(0) = 0, Z_j > t\}$ also occurs. When the queue is empty at time $t > 0$, the queue must be empty in the closed interval $[0, t]$, because c_j has departed at τ_j , which is the instant of the last previous departure taken to be the origin on the time axis by Burke, and because the next customer has not arrived as required by $0 < t < t_1$. According to Burke's theorem, namely, the interpretation of the mathematical results obtained by solving the equations, Z_j follows the exponential distribution of inter-arrival times with parameter λ . However, according to the chronological order of events experienced by customers, whenever $Q(t) = 0$,

$$Z_j = I_j + S_{j+1}.$$

Clearly, $I_j + S_{j+1}$ does not follow the distribution of inter-arrival times. Now consider $Q(t) > 0$, i.e., the queue is not empty at t , where t satisfies the inequality $t_1 < t < \tau_{j+1}$. Whenever $Q(t) > 0$,

$$Z_j = S_{j+1}$$

and S_{j+1} is not distributed as an inter-arrival time.

According to Lemma 1 and its corollary, for any j , there are only two kinds of realistic inter-departure times, T_1 and T_2 , which are already defined on $\Omega_1(j)$ and $\Omega_2(j)$, respectively (see Eq.(2.1) and Eq.(2.2) in Section 2); they cannot be expressed by any random variable on the whole sample space Ω . Such inter-departure times are terms of a sequence with two-fold randomness and cannot be characterized by a marginal distribution; their distributions cannot be determined by joint distributions of random vectors defined on Ω . As shown above, Burke's proof leads to a contradiction revealed in Section 2,

$$P_{T_1}(B) = P_{T_2}(B) = P_{Z_j}(B).$$

The contradiction is due to the false assumption underlying Burke's proof. Based on the assumption, $\tau_{j+1} - \tau_j$ is considered by Burke as a random variable Z_j on Ω with a marginal distribution obtained by treating $[Q(t), \tau_{j+1} - \tau_j]$ as a random vector, which leads to the following questionable interpretation of the mathematical results.

$$\sum_{k=0}^{\infty} \mathbb{P}[Q(t) = k, Z_j > t] = \mathbb{P}(Z_j > t) = e^{-\lambda t}.$$

Because Z_j is neither defined on Ω nor on any subset of Ω , it fails to describe realistic inter-departure times. The interpretation given by Burke ignores completely the dependence of departures on the state of the server. Such dependence is part of the constraints imposed by physical systems to be studied based on the queuing model.

3.2. Reich's Proof Based on Time-Reversible Markov Process

As a birth-death process in steady state (i.e., in statistical equilibrium), $Q(t)$ is a time-reversible Markov process. If instants on the time axis are ordered in the reversed direction, $Q(t)$ has a time-reversed counterpart $Q^*(t)$, which is also a birth-death process. If $Q(t)$ is used to model the size of a stable M/M/1 queue in steady state at time t , then "death" and "birth" represent "departure" and "arrival", respectively. Consequently, intervals between consecutive deaths are inter-departure intervals. Reich provided a different proof of Burke's theorem based on $Q^*(t)$ [6]. See also [7].

The proof based on $Q^*(t)$ is problematic. Although arrival instants when looking forwards in time constitute a Poisson process, the instant τ_j for any j at which $Q^*(t)$ increases by 1 cannot form a Poisson process, because $|\tau_j - \tau_{j+1}| = \tau_{j+1} - \tau_j$. If the birth-death process $Q(t)$ serves to model

the queue size of the M/M/1 system, then times between consecutive deaths and times between consecutive births do not necessarily follow the same distribution in whatever direction of time.

During an open interval between two consecutive deaths (i.e., departures from the M/M/1 queue), the state of $Q(t)$ may change m times, where $m \geq 0$. A change of $Q(t)$ during (τ_j, τ_{j+1}) is due to an arrival in (τ_j, τ_{j+1}) . If $m = 0$, then $Q(t)$ remains unchanged in (τ_j, τ_{j+1}) , which implies $Q(t) > 0$ for $\tau_j \leq t < \tau_{j+1}$, because at least one customer, c_{j+1} is still in the system, and because no customer arrives in the interval.

$$Q(t) = \begin{cases} k+2 & t < \tau_j \\ k+1 & \tau_j \leq t < \tau_{j+1} \\ k & t = \tau_{j+1}. \end{cases}$$

The inter-departure time $\tau_{j+1} - \tau_j = S_{j+1}$, because $Q(t)$ decreases by 1 if and only if one customer has been served.

If $m \geq 1$, then $Q(t)$ changes at least once during (τ_j, τ_{j+1}) . Let $t_i, i = 1, 2, \dots$ represent the instants at which $Q(t)$ changes; t_i is the i th arrival instant in the interval. Thus,

- either $Q(t) > 0$ for $\tau_j \leq t < \tau_{j+1}$,
- or $Q(t) = 0$ for $\tau_j \leq t < t_1 < \tau_{j+1}$, and $Q(t) > 0$ for $t_1 \leq t < \tau_{j+1}$.

$$Q(t) = \begin{cases} k+1 & t < \tau_j \\ k & \tau_j \leq t < t_1 \\ k+1 & t_1 \leq t < t_2 \\ k+2 & t_2 \leq t < t_3 \\ \dots & \dots \\ k+m & t_m \leq t < \tau_{j+1} \\ k+m-1 & t = \tau_{j+1}. \end{cases}$$

If $k > 0$, $\tau_{j+1} - \tau_j = S_{j+1}$ still holds. However, if $k = 0$, then

$$Q(t) = \begin{cases} 0 & \tau_j \leq t < t_1 \\ 1 & t_1 \leq t < \tau_{j+1} \end{cases}$$

where t_1 is the arrival instant of c_{j+1} , and $\tau_{j+1} - \tau_j$ consists of an idle time of the server (i.e., $t_1 - \tau_j = I_j$) and a service time (i.e., $\tau_{j+1} - t_1 = S_{j+1}$). Whatever value k takes, $\tau_{j+1} - \tau_j$ does not follow the distribution of inter-arrival times. Similarly, for $Q^*(t)$, if $m = 0$,

$$Q^*(t) = \begin{cases} k & t = \tau_{j+1} \\ k+1 & \tau_{j+1} < t \leq \tau_j \\ k+2 & t > \tau_j \end{cases}$$

and $|\tau_j - \tau_{j+1}| = S_{j+1}$. When $m \geq 1$,

$$Q^*(t) = \begin{cases} k+m-1 & t = \tau_{j+1} \\ k+m & \tau_{j+1} < t \leq t_m \\ \dots & \dots \\ k+2 & t_3 < t \leq t_2 \\ k+1 & t_2 < t \leq t_1 \\ k & t_1 < t \leq \tau_j \\ k+1 & t > \tau_j \end{cases}$$

and $|\tau_j - \tau_{j+1}| = S_{j+1}$ if $k > 0$; otherwise $|\tau_j - \tau_{j+1}| = I_j + S_{j+1}$. In steady state, times between consecutive departures from the M/M/1 queue always follow two different distributions, P_{T_1} and P_{T_2} , as indicated by Eq.(2.1) and Eq.(2.2) in Section 2. Clearly, the distribution of inter-arrival times is neither P_{T_1} nor P_{T_2} .

For a given t , $Q(t)$ is independent of future arrivals after t . Time reversibility is used here to argue for “ $Q(t)$ is also independent of past departures before t .” Based on an analogy between $Q(t)$ and $Q^*(t)$, the argument goes as follows: $Q(t)$ is independent of future arrivals after t ; departures are “arrivals” when looking backwards in time; because $Q(t)$ and $Q^*(t)$ are statistically identical, $Q(t)$ is independent of past departures before t .

However, the analogy between $Q(t)$ and $Q^*(t)$ is not appropriate, and the argument based on the analogy is problematic. Although $Q(t)$ is independent of arrivals after t , both arrivals and departures prior to t determine $Q(t)$: increase and decrease in $Q(t)$ are due to arrivals and departures before t , respectively. For a stable M/M/1 queue in steady state, $Q(t)$ depends on departures before t necessarily.

Any queuing model for solving problems in the real world must satisfy the constraints imposed by physical systems to be studied based on the model. The reversed process $Q^*(t)$ is constructed in pure mathematics without considering the chronological order of events experienced by customers and irrelevant to the departure process from the M/M/1 queue. By Corollary 1, $(\tau_{j+1} - \tau_j)_{j \geq 1}$ is a stochastic sequence with two-fold randomness; its terms cannot be characterized by a marginal distribution. Time reversibility cannot change $(\tau_{j+1} - \tau_j)_{j \geq 1}$ into a sequence of i.i.d. random variables on Ω . Thus, the arguments in queuing theory based on $Q^*(t)$ are questionable.

Questing the arguments in queuing theory based on $Q^*(t)$ is not to say “a birth-death process $Q(t)$ in statistical equilibrium has no time-reversed counterpart $Q^*(t)$.” The birth-death process $Q(t)$ is still time-reversible and has its time-reversed counterpart $Q^*(t)$. But $Q^*(t)$ is just not applicable to queuing models.

3.3. Experimental Results Concerning Counting Processes with Two-Fold Randomness

There are many experimental results obtained by simulation concerning departures from the M/M/1 queue. All of them are claimed to be in agreement with Burke’s theorem. However, as shown below, the experimental results are misinterpreted. Let $\mu < \infty$ represent the parameter of the service-time distribution. The mean inter-arrival time and the mean service time are $1/\lambda$ and $1/\mu$, respectively. Using the notations in Section 2, for all j ,

$$\mathbb{P}[\Omega_1(j)] = 1 - \rho$$

and

$$\mathbb{P}[\Omega_2(j)] = \rho$$

where $0 < \rho = \lambda/\mu < 1$. If $Q_j = 0$, then $\tau_{j+1} - \tau_j = T_1$, see Eq.(2.1), and its probability density function (pdf) is

$$f_{T_1}(t) = \frac{\lambda\mu}{\mu - \lambda}(e^{-\lambda t} - e^{-\mu t}).$$

If $Q_j > 0$, then $\tau_{j+1} - \tau_j = T_2$, see Eq.(2.2), and its pdf is $f_{T_2}(t) = \mu e^{-\mu t}$. Let K_t represent a time interval $(t, t + dt]$ of an infinitesimal length dt . Write

$$H_j = \{Q_j = 0, T_1 \in K_t\} \cup \{Q_j > 0, T_2 \in K_t\}.$$

For a stable M/M/1 queue in steady state, a simple calculation yields

$$\mathbb{P}(H_j) = \mathbb{P}(Q_j = 0, T_1 \in K_t) + \mathbb{P}(Q_j > 0, T_2 \in K_t) = \lambda e^{-\lambda t} dt. \quad (3.2)$$

The probability obtained by the calculation is indeed in agreement with the experimental results, but it cannot be interpreted as

$$\mathbb{P}(Q_j = 0, Z_j \in K_t) + \mathbb{P}(Q_j > 0, Z_j \in K_t) = \mathbb{P}(Z_j \in K_t) = \lambda e^{-\lambda t} dt. \quad (3.3)$$

In the above interpretation, Z_j is treated, incorrectly, as a random variable representing $\tau_{j+1} - \tau_j$ at each $\omega \in \Omega$.

By Lemma 1, $\tau_{j+1} - \tau_j$ is not a random variable defined on the whole sample space Ω , thus $(Q_j, \tau_{j+1} - \tau_j)$ is not a random vector. It is incorrect to interpret Eq.(3.2) as identical to Eq.(3.3), because $\{Z_j \in K_t\} \notin \mathcal{A}$ (see Section 2).

Simulation studies are based on the strong law of large numbers, and the sequence studied by simulation should consist of i.i.d. random variables on Ω . As shown in Section 2, $(\tau_{j+1} - \tau_j)_{j \geq 1}$ is a stochastic sequence with two-fold randomness; it can be randomly divided into two subsequences.

$$[\tau_{j+1}(\omega) - \tau_j(\omega)]_{j \geq 1} = [T_1(\omega)]_{j \in M_1(\omega)} \cup [T_2(\omega)]_{j \in M_2(\omega)}$$

which holds at any given $\omega \in \Omega$. Accordingly, $(H_j)_{j \geq 1}$ can also be divided randomly into two subsequences at the sample point.

$$[H_j(\omega)]_{j \geq 1} = [Q_j(\omega) = 0, T_1(\omega) \in K_t]_{j \in M_1(\omega)} \cup [Q_j(\omega) > 0, T_2(\omega) \in K_t]_{j \in M_2(\omega)}.$$

Consequently, it is impossible to construct a sequence of i.i.d. random variables based on $(H_j)_{j \geq 1}$, and it is not legitimate to apply the strong law to study $(H_j)_{j \geq 1}$. Nevertheless, if $\mathcal{I}(H_j)$, the indicator of H_j , is considered instead of H_j , then $\mathcal{I}(H_1), \mathcal{I}(H_2), \dots$ constitute a sequence of i.i.d. random variables all defined on Ω . By the strong law,

$$\lim_{n \rightarrow \infty} \frac{\sum_{j=1}^n \mathcal{I}(H_j)}{n} = \mathbb{E}[\mathcal{I}(H_1)]$$

with probability one, where

$$\mathbb{E}[\mathcal{I}(H_1)] = \mathbb{P}(Q_1 = 0, T_1 \in K_t) + \mathbb{P}(Q_1 > 0, T_2 \in K_t) = \lambda e^{-\lambda t} dt.$$

The above analysis not only explains why Eq.(3.2) is in agreement with the experimental results of simulation but also reveals the essential difference between Eq.(3.2) and Eq.(3.3); the interpretation given by Eq.(3.3) does not make sense in practice. To see this, consider an arbitrarily given j . Were the interpretation meaningful in applications, there would be a sample point $\omega \in \Omega$; at this ω , when c_j departs, a value would be taken by Z_j corresponding to both $Q_j(\omega) = 0$ and $Q_j(\omega) > 0$ simultaneously. In no sense can anything like this happen in the real world.

4. Two-Fold Randomness in Counting Processes: Networks of Queues

Consider two queues Q_1 and Q_2 in tandem, where Q_1 is an M/M/1 queue. Customers arrive first at Q_1 according to a Poisson process with a rate λ . After being served at Q_1 , they join Q_2 immediately; Q_2 also has an infinite waiting room. Service times of a customer spent at Q_1 and Q_2 are mutually independent, following exponential distributions with finite service rates μ_1 and μ_2 . In addition, service times at Q_2 are not only mutually independent but also independent of arrivals both at Q_1 and at Q_2 . If $\mu_1 > \lambda$, then Q_1 is stable. According to Burke's theorem, departures from Q_1 in steady state constitute a Poisson process with the rate λ . In the literature, Q_2 is claimed to be stable if $\mu_2 > \lambda$. However, because times between consecutive departures (i.e., $\tau_{j+1} - \tau_j$) from Q_1 are times between consecutive arrivals at Q_2 , and because $\tau_{j+1} - \tau_j$ cannot be described by any single, fixed random variable, the number of customers in Q_2 cannot be described by any single, fixed random variable either, even if Q_1 is stable and has reached its steady state. Burke's theorem allows the two queues in tandem to be treated as a Jackson network of queues [8]. The network consisting of such two queues in tandem is actually a counterexample to Jackson's theorem, which will be scrutinized next.

4.1. Jackson Networks of Queues and Jackson's theorem

Consider a Jackson network consisting of J single-server queues denoted by $Q_j, j = 1, 2, \dots, J$. At each queue, there is an infinite waiting room, and customers are served by a work-conserving server. The mean service time is $1/\mu_j$ at Q_j . By Jackson's theorem [8], these queues are stable if $\lambda_j < \mu_j$, where λ_j is the total arrival rate of customers at Q_j , determined by the following equations.

$$\lambda_j = \gamma_j + \sum_{i=1}^J \lambda_i P_{ij}, \quad i = 1, 2, \dots, J. \quad (4.1)$$

In Eq.(4.1), γ_j is the arrival rate of customers at Q_j from outside of the system, and P_{ij} is the probability for a customer to join Q_j immediately after leaving Q_i . So the arrival rate of customers at Q_j from Q_i is $\lambda_i P_{ij}$. All the proofs of Jackson's theorem, including Jackson's original proof and the proof with time reversibility, depend on the stability condition $\lambda_j < \mu_j$ obtained by solving Eq.(4.1).

According to Jackson's theorem and his interpretation (e.g. [8,9]), after the network has been in operation for an infinitely long time, each queue in the network behaved "as if"

- its size were a random variable characterized by a marginal distribution, and
- all such random variables were independent, possessing a joint distribution equal to the product of their marginal distributions.

However, the phrase "as if" in Jackson's interpretation makes it questionable, because it is inconsistent with the standard notion of independent random variables. For example, if the joint distribution of two random variables can be expressed as the product of their marginal distributions, the random variables are indeed independent, rather than behaved as if they were independent. The inconsistency is due to $\lambda_j < \mu_j, j = 1, 2, \dots, J$, the stability condition required by Jackson's theorem, which relies on an unjustified assumption (see below) and does not necessarily imply stability of *every* queue in the network. Consequently, the assumption makes Jackson's theorem irrelevant to physical systems modeled by Jackson networks of queues.

The unjustified assumption underlying Eq.(4.1) is this: times between successive departures from a stable queue in steady state follow a marginal distribution. This assumption is the basis to define the arrival rate at a queue for customers coming from inside of the system. By Lemma 1, for the network of the two queues in tandem, the assumption is problematic, unless the server at Q_1 has an infinite service capacity. Only in this unrealistic scenario, treating Q_2 as a stable M/M/1 queue isolated from Q_1 will not lead to the inconsistency. Because service capacities in the real world must be finite, it is illegitimate to use λ , the arrival rate at the first queue, to characterize the arrivals at the second queue. As terms of a stochastic sequence with two-fold randomness, $\tau_{j+1} - \tau_j$ do not have a marginal distribution. Time reversibility cannot change this fact or explain away the inconsistency mentioned above, see also Subsection 3.2.

In general, so long as inter-arrival times between customers at a queue are times between consecutive departures from another queue, any single, fixed random variable cannot describe such inter-arrival times. Consequently, any single, fixed random vector cannot describe the behavior of a Jackson network of queues, regardless of whether the structure of the network is simple or complex. In a Jackson network, with or without feedback paths, at least one queue is not stable. That is, statistical equilibrium with respect to the numbers of customers in all the queues in the network as a whole does not exist. Therefore, Jackson's theorem is indeed questionable, and no Jackson network is stable, although the solution of Eq.(4.1) is not difficult to find.

By definition [4], if the number of customers in a queue remains finite after the queue has been in operation for an infinitely long time, the queue is *sub-stable*. A stable queue is of course sub-stable. But a sub-stable queue may not necessarily be stable. If a queue is sub-stable but not stable, the queue is *properly sub-stable*. If a queue is properly sub-stable, the number of customers in the queue is always finite, but its distribution will not converge to a limit. That is, the behavior of the queue cannot be described by any single, fixed random variable. The meaning of "not stable" is not "unstable"; the

latter means “not sub-stable”. The number of customers in an unstable queue will become infinitely large as time approaches infinity. As shown above, in a Jackson network consisting of two queues at least, a properly sub-stable queue is mistaken for a stable queue.

4.2. Generalization

It is possible to generalize the results concerning Jackson networks of queues. Consider a system of work-conserving, single-server queues in tandem. Each queue has a finite service capacity and an infinite waiting room. At each queue, service times are generally distributed, mutually independent, and independent of its arrivals. The first queue is a stable GI/GI/1 queue in steady state. All customers arrive from outside of the system at the first queue and leave the system from the last queue after being served there. A queue in the system is called a downstream queue, if it is not the first queue or the last queue. After receiving service at the first queue or a downstream queue, a customer goes immediately to the next queue. A queue in the system is merely properly sub-stable rather than stable, if it is not the first queue. Consequently, this system of queues in tandem as a whole is not stable, in the sense that its behavior cannot be described by any single, fixed random vector. For the downstream queues and the last queue, the two different notions, proper sub-stability and stability, are confused in the literature; the confusion leads to mistaking properly sub-stable queues for stable queues. Such proper sub-stability due to the departure processes with two-fold randomness is entirely ignored in the existing literature.

Other possible generalization is to apply the above analysis to queuing networks with more general topological structures. For example, any queue in a network may have multiple work-conserving servers with finite service capacities; its waiting room may not necessarily be infinite. External arrivals at the queue may not necessarily form a renewal process. Service times may follow different distributions at different queues in the network; external arrivals and service times can also be dependent.

5. Conclusion

Phenomena with multiple randomness appear naturally in random experiments concerning practical applications of sciences and engineering disciplines. Many scientists and engineers use counting processes to study random phenomena in the real world. However, counting processes with multiple randomness, which differ essentially from known stochastic processes in the existing literature, are all mistaken for usual counting processes, because correctly derived mathematical results are used to interpret, incorrectly, the corresponding experimental results, leading to questionable theorems in practice. This article introduces general results concerning multiple randomness in counting processes and provides specific examples in queuing theory to illustrate the properties of two-fold randomness. In addition to the examples in queuing theory, more counting processes with m -fold randomness ($m \geq 2$) may be found in other application fields of sciences and engineering disciplines and in other stochastic processes used by scientists and engineers, which may help them to explain weird phenomena and solve puzzling problems in the real world.

Funding: This research received no funds or grants.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

References

1. Kloska1, S.M.; Pałczyński, K.; Marciniak, T.; Talaśka, T.; Miller, M.; Wysocki, B.J.; Davis, P.; Wysocki, T.A. Queueing theory model of pentose phosphate pathway. *Scientific Reports* **2022**. <https://doi.org/10.1038/s41598-022-08463-y>.
2. Buzna, L.; Carvalho, R. Controlling congestion on complex networks: fairness, efficiency and network structure. *Scientific Reports* **2017**. <https://doi.org/10.1038/s41598-017-09524-3>.
3. Baskett, F.; Chandy, K.M.; Muntz, R.R.; Palacios, F.G. Open, closed, and mixed networks of queues with different classes of customers. *Journal of the ACM* **1975**, 22, 248–260. <https://doi.org/10.1145/321879.321887>.
4. Loynes, R.M. The stability of a queue with non-independent inter-arrival and service times. *Proc. Camb. Philos. Soc.* **1962**, 58, 497–520. <https://doi.org/10.1017/S0305004100036781>.
5. Burke, P.J. The output of a queueing system. *Operations Research* **1956**, 4, 699–704. <https://doi.org/10.1287/OPRE.4.6.699>.
6. Reich, E. Waiting times when queues are in tandem. *Ann. Math. Statist.* **1957**, 28, 768–773. <https://doi.org/10.1214/AOMS/1177706889>.
7. Kelly, F.P. *Reversibility and Stochastic Networks*; John Wiley & Sons: New York, 1979.
8. Jackson, J.R. Networks of waiting lines. *Operations Research* **1957**, 5, 518–521. <https://doi.org/10.1287/OPRE.5.4.518>.
9. Jackson, J.R. Job-shop like queueing systems. *Management Science* **1963**, 10, 131–142. <https://doi.org/10.1287/MNSC.1040.0268>.