IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 July 2024, accepted 22 August 2024, date of publication 28 August 2024, date of current version 10 September 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3451181

RESEARCH ARTICLE

Excluded Volume Effect in General Distributions

TAKAKO HOSHIYAMA, (Graduate Student Member, IEEE), AND KATSUHIRO NISHINARI
Department of Advanced Interdisciplinary Studies, School of Engineering, The University of Tokyo, Meguro-ku, Tokyo 153-8904, Japan
Department of Aeronautics and Astronautics, School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

Corresponding author: Takako Hoshiyama (thoshiyama@g.ecc.u-tokyo.ac.jp)

ABSTRACT In queuing systems, the dynamics of the excluded volume effect, which describes the movement of objects with a physical volume, are generally not considered. Recent studies have presented a theoretical formulation that incorporates the excluded volume effect to describe the delayed dynamics in M/M/1 queuing systems. We attempted to extend these results to the GI/G/1 queuing systems. Although the GI/G/1 queuing model is a stochastic process that cannot be completely described by a Markov chain, recent studies have indicated that it can be solved numerically using matrix geometry methods in discrete time under certain conditions. We propose a theoretical and analytical methodology in addition to approximation formulas for a discrete-time GI/G/1 queuing model that incorporates the excluded volume effect by utilizing the matrix-geometric method. The approximation methods were validated under single-server conditions for practical applications. Furthermore, using a theoretical perspective based on the aggregation method, we conducted numerical experiments on multiple servers to evaluate the performance of the discrete-time GI/PH/c queuing model. This approach adapted the excluded volume effect to the geometric matrix method. After testing it with a hyper-gamma distribution, we observed practical agreement, albeit to a limited extent.

INDEX TERMS Aggregation, ASEP, excluded volume effect, GI/G/1, GI/PH/c, matrix-geometric method.

I. INTRODUCTION

Since its discovery, the asymmetric simple exclusion process (ASEP) model has been associated with various phenomena in the physical and chemical sciences [1], [2], [3]. It is a multiparticle system moving in a one-dimensional (1D) lattice with an excluded volume rule, such as the congestion of ribosomes on messenger ribonucleic acids. ASEP was first introduced in 1968 to explain protein synthesis [4]. This yields exact solutions for these systems [3], [5], [6], [7], [8], [9], [10]. However, the queuing theory does not consider the excluded volume effect. The ASEP is more likely to occur when the waiting situation is prominent in a queuing system [6], [11]. The divergence of the system in ASEP is characterized when the utilization rates of the input and output are less than one, whereas the collapse of the system in the queuing theory occurs when the utilization rate reaches one. Some studies have shown that the characteristics of this behavior in M/M/1 queuing models can be explained by discrete time, update, and excluded volume effects in the discrete lattice space [3], [6]. ASEP has been extended to traffic flow modeling, and the production network is one of its applications [12]. In this study, ASEP theory can be extended to an open GI/G/1 queuing model as a novel approach. We also propose formulas for approximations adapted to a limited region of the queuing model using the excluded volume effect. Ordinary queuing theory does not consider the concept of volume formation in the queue; therefore, it is inadequate for some problems in which the size of the objects has an effect.

The ASEP model, which incorporates the excluded volume effect, is more prone to congestion than the normal queuing model [6], [11]. The ratio of the probability of n particles entering a system to that of a particle leaving the system is known as the density $\rho(n)$, and the extremely high congestion in queuing theory does not require the concept of volume. This is known as a critical point, occurring at $\rho = 1$ in ordinary queues, where the volume of the objects is not considered. Some studies are also underway to incorporate the excluded volume effect into queuing theory. This is expected to provide a result that more accurately represents the behavior of physical systems. Yanagisawa et al. introduced the excluded volume effect into an M/M/1 queuing

The associate editor coordinating the review of this manuscript and approving it for publication was Engang Tian.

model that derived exact solutions from the master equation and demonstrated that the critical point was lowered [3], [5] [10]. However, this model is currently limited by the scope of M/M/1 and has a limited range of applications.

Several queuing theories have been published in which the probability distributions of arrivals and departures are extended to general distributions. However, owing to the complexity of theories and difficulty in calculation, approximation methods have been developed to predict reasonable results [11], [13], [14], [15]. Among the numerous ideas, the formula proposed by Sakasegawa [16], which can be applied to GI/G/1 and GI/G/s, has been evaluated as being practical. The formula in [16] that has been used in this study has been reported as a good fit in the region $\left\{c_a^2, c_s^2\right\} \leq 1.0^2$ [17]. The matrix geometric method (MGM) uses the M/M/1 queuing model. The MGM was developed using Markov chain analysis. Later, Evance et al. proposed an algorithm for verification and contributed to the study of the extension of the types for the queuing model types. Wallace et al. developed an MGM by employing a Markov chain in a traditional M/M/1. This model features a continuous-parameter Markov chain with block Jacobi generators similar to the GI/M/1 type. The stability conditions within the MGM were established by Neuts et al. (1975). This also involves a continuous-parameter Markov chain with block Jacobi generators resembling the GI/M/1 type [18]. Moreover, this condition uses the properties of a Markov chain with an infinite block structure of the quasi-birth–death process and continuous parameters with transition matrices to enable us to obtain the precise analysis of GI/M/1 and M/G/1. This method employed the same stability conditions as the evaluation indicator. Subsequently, Alfa and Li. proposed a method capable of calculating GI/G/1 using the same calculation method as the discrete-time PH/PH/1; hereinafter, this model is described as the Alfa model [19], [20], [21]. In this study, we propose using this method to extend the adaptation range of the excluded volume effect to the approximation method proposed by Sakasegawa, and then apply the Alfa model [22]. We also extend the model by adapting the behavior of the totally asymmetric simple exclusion process (TASEP) to the GI/G/1. In addition, we associated the aggregation method [22] with the Alfa model (hereinafter referred to as the aggregated Alfa model) and evaluated the performance of the TASEP when adapted to GI/PH/c with a hyper-gamma distribution.

II. THEORETICAL PREPARATION AND APPLICATIONS

A. QUEUE LENGTH L_q IN THE QUEUING THEORY

Queuing theory defines the events of arrival and processing in a system as the average time interval (expected value) represented by the inverse of the average value of the time intervals λ and μ, respectively (Fig. 1). Taking the M/M/1 queuing model as an example, because the time intervals of arrival and processing are stochastic variables X and Y, they should follow an exponential distribution whose probability density function is given by $p(x) = \lambda exp(-\lambda x)$ *and* $p(y) = \mu\ exp(-\mu y)$. The mean values of these random variables were $E(X) = 1/\lambda$ *and* $E(Y) = 1/\mu$, respectively. The number of particles remaining in the system, excluding those processed in the server, is known as the queue length $L_q = \rho^2/(1-\rho)$, as $\rho = \lambda/\mu$. The denominator of this equation is zero at $\rho = 1$; therefore, L_q diverges at $\rho = 1$. This implies that when the inverse of the mean of the arrival time intervals λ is equal to the inverse of the mean of the processing time intervals μ, it exceeds the processing capacity of the system, which is known as the critical point [23], [24].

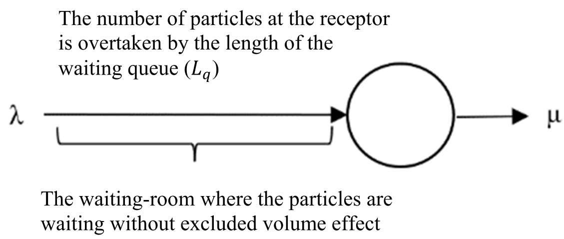

FIGURE 1. Relation between the waiting room and the queue length L_q in the queuing theory.

B. ASEP MODEL AND EXCLUDED VOLUME EFFECT

The ASEP model strictly regulates the events at the critical point of the excluded volume effect. The model incorporates the concept of transition probabilities into the deterministic cellular automaton model and is a system with the behavioral property of transitioning with probability $p\ (0 \leq p \leq 1)$, when the front space is vacant. The transition probability for returning is denoted by $q\ (0 \leq q \leq 1)$. In particular, when both probabilities become $p = 1$ and $q = 0$, this corresponds to Rule 184 of the cellular automaton, which is equivalent to the TASEP system. The particles moving in the 1D discrete lattice of the TASEP system are shown in Fig. 2. The figure shows that the excluded volume effect restricted the flow rate of the particles.

Inflows and outflows in the queuing model are defined as the inverse of the mean of the time interval λ for inflows and μ for outflows. The inflow and outflow events in the TASEP model are specified as one of the boundary conditions; that is, when the first site on the left-hand side of the open space is empty, it enters with a probability α from the first site on the left-hand side. In addition, when the first leftmost site was occupied by another particle, no new particles could enter. In other words, the ASEP model is a call-loss system, with γ corresponding to its rate. If the neighboring site is empty, the particle jumps to the right or left site with a certain probability p or q. If the particle eventually moves to the rightmost site, it exits the system with a probability β. The particle may not be able to exit if an obstacle is present outside the exit when it attempts to move forward with probability δ. The model in this study employed an open TASEP system and assumed that $q = \gamma = \delta = 0$. When either the transition probability p or q is zero, it is referred to as TASEP. Fig. 2 shows the excluded volume effect in the case of TASEP with $p = 1$, where the backward particle cannot jump if the forward neighboring site is occupied, and the "excluded volume effect" is marked with "×." In the same figure, the middle image illustrates an event without an excluded volume effect. We consider a 1D discrete lattice of a discrete-time system with inflow probability α, transition probability p, and outflow probability β corresponding to the excluded volume effect.

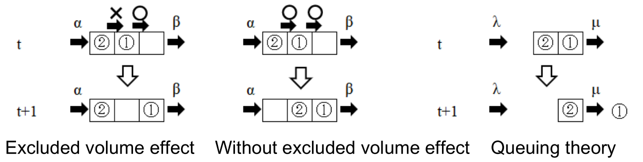

FIGURE 2. **($p = 1, q = 0$) Diagram of the TASEP and queuing theory.**

Although several options exist for the update method, we select the rule of parallel updates for discrete-time systems [3], [5] [25] [26]. In a queue in steady state, no more than two free lattice spaces can occur (hereinafter denoted as the 2^n state) [3]. The particle behaviors of the state transition in the single-receptor TASEP model is defined as follows:

Time, $t{:}\{t_0, \cdots, t_n\}, n \in \mathbb{N}_0$

Cells, $Cell{:}\{cell_0, \cdots, cell_n\}, n \in \mathbb{N}_0$

$cell_0$: *System entrance cell with excluded volume effect*

$cell_n$: *System exits and receptor cell with excluded volume effect*

Particles, $P{:}\{p_0, \cdots, p_n\}, n \in N_0$

States *with behavioral rules: moving, waiting, processing, arriving / leaving 1D discrete lattice space, where:*

$p_s = $ *Particle which is arriving at the cell*, $cell_0$

$p_m = $ *Particle which can move from the cell*, $cell_1$ *to* $cell_{n-1}$ *in the time*, $t + t_1$ *with probability* 1

$p_w = $ *Particle which can wait on a cell*, $cell_1$ *to* $cell_{n-1}$

$p_{pro} = $ *Particle which is processing in the cell*, $cell_n$

$p_d = $ *Particle which is leaving from the cell*, $cell_n$

The detailed dynamics of the particles in a 1D discrete lattice are shown in Fig. 2. For a detailed description of this method, refer to [2]. The relationship between the flow rate J and the density layer of the open TASEP model in the phase diagram is presented as follows:

$$J(\alpha, \beta, p) = \begin{cases} \alpha \dfrac{p-\alpha}{p-\alpha^2} \cdots LD(Lower - Density)\ regime, \\ \beta \dfrac{p-\beta}{p-\beta^2} \cdots HD(High - Density)\ regime, \\ \dfrac{1-\sqrt{1-p}}{2} \cdots MC(Maximal - Current)\ regime. \end{cases}$$

Fig. 4 shows the curve representing the critical point of the M/M/1 queuing model with the excluded volume effect. This demonstrates the dynamics of the boundary line between the lower-density (LD) and high-density (HD) phases, along with the region between the maximal current (MC) and LD (or MC and HD) phases, when the state of the critical point is plotted

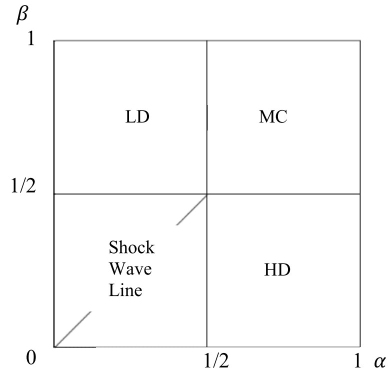

FIGURE 3. **($p = 1$) Phase diagram of open TASEP.**

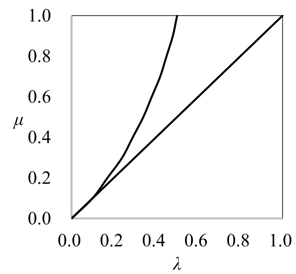

FIGURE 4. **($p = 1$) Comparison between the critical point state in the M/M/1 model incorporating the excluded volume effect and the M/M/1 queuing model: (The curve is the critical point of the M/M/1 model with the excluded volume effect, plotted by (1), and the line is the critical point of the queuing theory, plotted by $\rho = \lambda/\mu$. The region around the upper left of the figure corresponds to the "Convergence" events that occur around the critical line of each model, and the region around the lower right of the figure corresponds to the "Divergence" events.).**

with $\alpha = 1$ and $0 \leq \beta \leq 1$ in Fig. 3. The boundary between the LD and HD phases corresponds to the shockwave line, which exhibits dynamics of random-walk dynamics [2].

III. INTEGRATION OF QUEUING THEORY AND ASEP MODEL

A. COMPARISON OF THE CRITICAL POINT STATE IN THE M/M/1 MODEL INCORPORATING THE EXCLUDED VOLUME EFFECT AND M/M/1 QUEUING MODEL

In recent years, the same model has been used for rigorous analysis of traffic congestion phenomena, such as pedestrians and vehicles as self-driving particles [3], [10]. The concepts of distance, acceleration and deceleration events, and spaces between pedestrians are rigorously understood. Subsequently, the congestion events due to their delays were quantitatively analyzed as flow rates to clarify behavioral dynamics. In [3], the limit of the master equation was derived from a combination of M/M/1 queuing theory and the excluded volume effect. A remarkable result was that $L_q = \tilde{\rho}/(1-\tilde{\rho})$, where $\tilde{\rho} = \lambda/(1-\lambda)\,\mu$. This outcome leads to the critical point shown in (1), which can be interpreted as the critical point reached when $\tilde{\rho} < 1$. This also indicates that the excluded volume effect of the TASEP shifts the critical point toward a lower arrival probability.

The derivation of (1) is presented in the Appendix. For the discrete-time open M/M/1 queuing model, μ and λ in (1) are treated as probabilities. However, μ and λ in this study were interpreted as the inverse of the mean of the time intervals in discrete time, with $t_{\mathbb{N}} = 1$.

$$\lambda_{cr} = \frac{\mu}{1+\mu}. \qquad (\lambda_{cr} > \lambda = \mu) \tag{1}$$

A plot comparing the critical point state (curve) from (1) with that (straight line) from the conventional M/M/1 queuing model is shown in Fig. 4, which has the same dynamics as in [5].

B. SAKASEGAWA APPROXIMATION

We now consider an extension of the GI/G/1 queuing model and discuss the approximation formula presented by Sakasegawa [16]. This formula is an extension of that derived in an earlier study by Page [17]. Sakasegawa also provided approximation formulas for both the GI/G/1 and GI/G/s models. It was devised by a numerical method utilizing simulation without numerical tables, and it inherits the approximation formula that fits well in the region of $\{c_a^2, c_s^2\} \leq 1.0^2$ [17]. The GI/G/1 queuing model adopted in this approximation assumes that the distribution of the inverse of the mean of the arrival and service time intervals with infinite queues is an exponential distribution of the no-memory processes. If the packet arrivals follow a stationary Poisson process, and the service times follow independent and identical exponential distributions, the single-server model agrees well with the M/M/1 queuing model. Here, we introduce (2) and (3) as approximations of [16]: $L_{q(\mathrm{GI/G/1})}$ denotes the average queue length in GI/G/1, and $\{c_a^2, c_s^2\}$ are the coefficients of variation (square of the variance divided by the mean) of the arrival and processing random variables, respectively. The critical point of this equation is when $\rho = 1$, because the excluded volume effect is not considered. However, factor $(c_a^2 + c_s^2)/2$ is assumed to represent the specific behavior of the general distribution, which is consistent with the M/M/1 results, as expected when $\{c_a^2, c_s^2\} = 1.0^2$.

$$L_{q(GI/G/1)} \cong \frac{c_a^2 + c_s^2}{2} \cdot \frac{\rho^2}{1-\rho}. \tag{2}$$

$$E\left(W_{GI/G/1}\right) \cong \frac{c_a^2 + c_s^2}{2} \cdot \frac{\rho}{1-\rho}\tau. \tag{3}$$

$E\left(\mathrm{W_{GI/G/1}}\right)$ denotes the expected value of the waiting time obtained from the approximation formula in GI/G/1 and τ is the processing time of a node with the same meaning as $1/\mu$. The utilization rate ρ in the queuing system in (2) and (3) is defined as $\rho = \lambda/\mu$.

IV. NEW INTERPRETATION OF THE EXCLUDED VOLUME EFFECT IN THE TASEP MODEL

A. FLOW AND VELOCITY IN STEADY STATE

We adapted the critical point relation $\bar{\rho}$ of the TASEP system, which excludes the volume effect of (4), to ρ in the same equation. In previous studies, [3], [10] proposed a combination of TASEP theory with M/M/1 and used $\hat{\rho} = \lambda(1+\mu)/\mu$. This implies that the excluded volume effect is related to the relation of modification, $\mu \to \mu/(1+\mu)$. However, as $\mu \to \mu/(1+\mu)$, it does not include the condition that $\mathrm{c_a^2}, \mathrm{c_s^2} \neq 1.0^2$; Thus, adapting it to the distribution of GI/G/1 is not appropriate. Instead, we consider the physical implications of the TASEP and incorporate the excluded volume effect because it causes a delay as the reciprocal of the average processing time intervals. In other words, the expected value of the processing time interval is "$\mu \to \mu$ (the probability of a particle moving forward)" with the excluded volume effect adapted to the GI/G/1 queuing model. The probability of a particle moving forward is estimated from the velocity in the TASEP model. The same equation also corresponds to a change in the critical point owing to a change in the mean value by the coefficient of variation to accommodate a general distribution, which also includes the excluded volume effect of the TASEP. This implies that the probability of a particle being present in a given cell n in a 1D discrete lattice is obtained by the density $\rho(n)$ [2]. The probability that no particle exists in cell $n+1$ can be expressed as $1 - \rho(n+1)$. If the transition probability is p_{hop} and the total number of particles in the 1D discrete lattice is N, the flow rate is obtained as follows:

$$Q = (1/N) \sum_{n=0}^{N-1} p_{hop}\rho(n)\{1 - \rho(n+1)\}$$

where the first term $(1/N)$ can be interpreted as the total number of particles N because it is a constant ρ_{mean} independent of the probability of particles remaining in n-th cell in a steady state [2]. In other words, the flow rate $Q_{const.}$ under steady-state conditions is $Q_{const.} = p_{hop}\rho_{mean}(1 - \rho_{mean})$. If the average velocity of particle transition in a 1D lattice is v_{mean}, and from $Q_{const.} = \rho_{mean}v_{mean}$, we obtain the following:

$$v_{mean} = p_{hop}(1 - \rho_{mean}).$$

B. INTERPRETATION OF VELOCITY AS PROBABILITY OF PARTICLE ADVANCEMENT IN THE SYSTEM

The TASEP system in the discrete model is dimensionless, because the smallest unit of a cell is 1. This spatial definition is not described in units of velocity [m/s], as in hydrodynamics. However, the system checks the state of the cell ahead of each epoch and determines whether or not to proceed based on the transition probability. An event that moves forward by one cell per unit of time has one as the velocity. In the TASEP model, $v_{mean} = p_{hop}(1 - \rho_{mean})$ is the only event that advances because of the excluded volume effect. Velocity v_{mean} is an event in which the forward cell of a particle in a 1D lattice is vacant, and the next cell is not vacant. This velocity is essentially the probability that the forward cell of a particle is free and that an event occurs in which the particle transitions the next time. This event indicates a delay in velocity owing to the excluded volume effect, which is expressed as a probability. This probability is defined as p_e

to be verified using the MGM. In addition, this probability is applied to the inverse of the mean value of the time interval of the processing time of the queuing model μ. The inverse of the mean of the processing time interval after adapting p_e is defined as μ', which is expressed in (4) as follows.

$$\bar{\rho} = \frac{\lambda c_a^2}{\mu' c_s^2}. \quad \left(\mu' = \mu p_e, p_e = p_{hop}\left(1 - \rho_{mean}\right)\right). \tag{4}$$

When the transition probabilities were $p_{hop} = 1, q = 0,\ and\ \rho_{mean} = 0.5$, the equivalent amount was $p_e = 0.5$. The steady-state TASEP system with forward probability $p_{hop} = 1\ and\ q = 0$ comprises $2L$ number of cells, where the number of empty cells and that of particles in system L are exactly half each. This is based on the configurational characteristics of the particles in a 1D lattice, with the relation $J = \left(\alpha - J^{out}\right) = \rho v$. The relation $Q_{const.} = \rho_{mean} v_{mean}$ represents the dynamics of the total system flow in the 1D discrete lattice of the TASEP. J is expressed as the difference in the outgoing flow J^{out} from the inverse of the mean value of the intrusion time interval α, which is the "net flow." When the left-hand side is zero, the arrival and departure are equal, implying $v = 0\ and\ \alpha \to \alpha_c$ (critical point) in a 1D lattice. From (21) in [7], $v = \alpha - \beta + \alpha\beta$ becomes $v = 2\alpha - 1$ when $\beta = 1$, and $v = 0$ becomes a critical point when $\alpha = 1/2$ and $\alpha = J^{out}$. In this study, these physical quantities in a 1D discrete lattice space were extended to the GI/G/1 queuing model; therefore, a gamma distribution was used for the mean values of the time intervals between arrivals and departures. For the gamma distribution, when the mean of the time interval between arrivals $1/\lambda \mathrm{c}_\mathrm{a}^2$ and departures $1/\mu \mathrm{c}_\mathrm{s}^2$ was equal, $\left\{c_a^2, c_s^2\right\} = 1.0^2$ and $\bar{\rho} = \lambda \mathrm{c}_\mathrm{a}^2/\mu \mathrm{c}_\mathrm{s}^2 = 1$ was the equivalent critical point. By replacing the inverse of the average time interval on the processing side of the queuing theory, μ with $\mu' = \mu p_e$, we provide the queue length and waiting time as (5) and (6). This is an approximation of the GI/G/1 queuing model that adapts the relationship in (4), including the excluded volume effect.

$$L_{q(GI/G/1)_{ASEP}} \cong \frac{c_a^2 + c_s^2}{2} \cdot \frac{\bar{\rho}^2}{1 - \bar{\rho}}. \tag{5}$$

$$E\left(W_{GI/G/1}\right)_{ASEP} \cong \frac{c_a^2 + c_s^2}{2} \cdot \frac{\bar{\rho}}{1 - \bar{\rho}} \tau. \tag{6}$$

These are the new approximations proposed in this study. All the phenomena in which objects of real physical size move contain an excluded volume effect. Therefore, the proposed approximation formula is expected to be useful for a more realistic understanding of object behavior.

V. INTRODUCTION OF EXCLUDED VOLUME EFFECT INTO THE ALFA MODEL

We extended the residual time vector of the processing time interval of the Alfa model [22] using MGM by adapting a delay probability corresponding to the excluded volume effect. In addition, a model that adapted the aggregation method to the same model was used. An overview of the MGM and aggregation method is provided in the Appendix (MGM and Alfa model) in this paper. This is discussed in detail below.

A. MODIFIED MATRIX S

To assess the validity of our proposed approximation, we introduce the excluded volume effect as the forward probability of a particle into the Alfa model, in which the forward direction of a particle on the processing side is represented by the matrix S. If the matrix S before the operation is identified as S_0 and the matrix after the operation as S, appropriate modifications are performed, as shown in (7). Matrix P_S is a time transition matrix operating on a row vector $\boldsymbol{\beta}_j = \left[\beta_1, \beta_2, \cdots, \beta_{n_s}\right]$, which represents the time remaining time until the particle departs from the system.

Probability p_e expresses the amount of delay in the transition of the TASEP system with the excluded volume effect. In general, the unit matrix I has infinite dimensions, that is, $I = I_{\infty \times \infty}$. Matrix S_0 must operate before probability p_e. This probability is the same as p_e in (4).

$$\begin{aligned} S &= p_e S_0 + (1 - p_e) I, \\ P_S &= \begin{bmatrix} \mathbf{1} & 0 \\ \boldsymbol{s} & S \end{bmatrix}, \\ S_0 &= \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \end{aligned} \tag{7}$$

The quantity at this critical point is known from previous studies to be $p_e = 0.5$ when the coefficients of variation $\left\{c_a^2, c_s^2\right\} = 1.0^2$ and the transition probability is 1. When $p_e = 0$, the state transition probability indicates that the process has not proceeded. However, when $p_e = 1$, the result was consistent with that of the Alfa model. Adapting (7) to the second equation of matrix S as the transition probability vector $\boldsymbol{s}$, which represents the state change in the remaining processing time, it can be transformed as follows, resulting in (8).

$$\begin{aligned} \boldsymbol{s} &= 1 - S\mathbf{1} \\ &= (I - S)\,\mathbf{1} \\ &= \{I - p_e S_0 - (1 - p_e) I\}\,\mathbf{1} \\ &= (-p_e S_0 + p_e I)\,\mathbf{1} \\ &= p_e (I - S_0)\,\mathbf{1}. \end{aligned} \tag{8}$$

Vector $\mathbf{1}$ in (8) is a vector whose elements are all 1's, and the number of elements is the same as the number of columns in matrix S, therefore, product $S\mathbf{1}$ is the sum of all the column vectors of matrix S and represents the sum of the probabilities of all state changes represented by matrix S; Thus, subtracting it from 1 represents the probability of a state change not represented by S. In other words, vector $\boldsymbol{s}$ represents the probability of all changes in state other than those represented by S. When matrix S, representing the change in the state of the remaining processing time of a particle, is changed by $p_e = 0.5$, the probability element p_e of the longitudinal vector element $\mathbf{s}$ of the transition matrix P_S moves to a state with less waiting time and a probability 0.5. Probability $1 - p_e$ of the same longitudinal vector element remains unchanged with probability $1-(0.5)$, indicating that the overall transition probability is reduced by a factor of p_e. The modification of the processing-time evolution introduces an effect such that the time progress after arrival becomes p_e times (or $1/p_e$ times) relative to the entire period.

The above equations were adapted to the MGM algorithm to compute the model evaluation as follows:

i. Initialization
ii. Service and arrival distribution settings
iii. Build state transition matrices S, T
iv. Iterate and determine convergence of matrix R
v. Eigenvalue analysis of the transition matrix P that is stationary
vi. Obtaining the steady state vector
vii. Evaluating system performance (e.g., computing L_q)

VI. NUMERICAL VARIDATIONS AND LIMITATION

In this section, we summarize the validation results used to evaluate the approximation proposed in (5). The Alfa model, adapted to the excluded volume effect, was used for the validation. Sakasegawa's approximation (2) for parameters $\{c_a^2, c_s^2\}$ less than 1.0^2 is comprehensively fitted with GI/G/1 and exhibits high agreement [16]. It is also known to consider the dynamics as an M/M/1 queuing system when the parameters $\{c_a^2, c_s^2\}$ in approximation (2) are equal to 1.0^2. The dynamics of the M/M/1 queuing system adapting to the excluded volume effect have also been understood in previous studies [3]. Thus, we verified the above properties using the proposed approximation (5). Additionally, the GI/PH/c queuing model was tested using the aggregation method with a hyper-gamma distribution. Practical results are obtained for the GI/G/1 queuing system by adapting it to the excluded volume effect.

A. VALIDATION WITH M/M/1 QUEUING SYSTEM WITH EXCLUDED VOLUME EFFECT

Fig. 5 shows the validation results of the M/M/1 queuing system with the excluded volume effect and the Alfa model with the adapted excluded volume effect.

These results exhibit a high level of agreement between the proposed approximation formula (5) and the Alfa model with the adapted excluded volume effect under the condition of $\{c_a^2, c_s^2\} = 1.0^2$. The mean absolute value of the percentage of the overall difference was 4.70360%.

B. DETAILED VERIFICATION BY GI/G/1 QUEUING SYSTEM WITH / WITHOUT EXCLUDED VOLUME EFFECT

We plot the results of the validation under limited conditions and the calculation of the proposed approximation (5) for

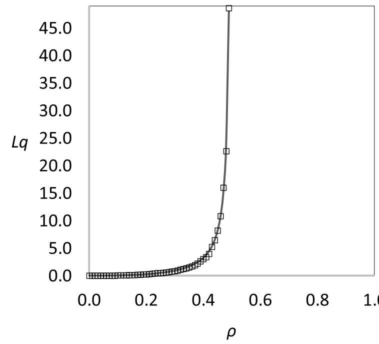

FIGURE 5. Comparison of $L_{q(\mathrm{M/M/1})\mathrm{ASEP}}$ between the approximation (5) and Alfa model with the adapted excluded volume effect under the conditions $\{c_a^2, c_s^2\} = 1.0^2, dt = 0.1, p_e = 0.5$ (Vertical axis: $L_{q(\mathrm{M/M/1})\mathrm{ASEP}}$, horizontal axis: ρ, solid curve: from (5), square dots: Alfa model with the adapted excluded volume effect).

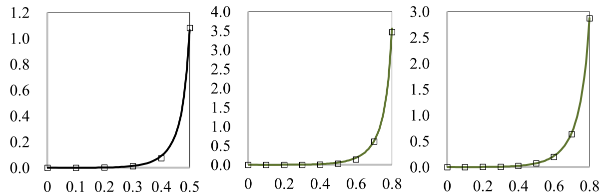

FIGURE 6. Comparison of $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$ between the proposed approximation formula (5) and the Alfa model under the conditions $\{c_a^2, c_s^2\} < 1.0^2, p_e = 0.5$. (From left figure to right: ($c_s^2 = 0.3^2$ $\rho = 0.16$), ($c_s^2 = 0.6^2 \rho = 0.25$), and ($c_s^2 = 0.8^2 \rho = 0.42$), respectively. Vertical axis: $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$, horizontal axis: c_a^2, both variables, solid curve: from (5), square dots: Alfa model).

the Alfa model adapted to the excluded volume effect in Fig. 6. For comparison with existing models, the performance evaluation against Sakasegawa's approximation (2) for the Alfa model without adaptation to the excluded volume effect (inthiscase :$p_e = 1.0$) is also listed in Table 3 in the Appendix. Note that some of the results of the performance evaluation of the Alfa model without the excluded volume effect using Sakasegawa's approximation (2) have been in published [22], [27]. All the detailed results are summarized in the Appendix of this paper in Table 3 Performance Evaluations of (2) and (5) (case: $\{c_a^2, c_s^2\} \leq 1.0^2, p_e = \{0.5, 1.0\}, dt = 0.1$).

1) CASE: $\{c_a^2, c_s^2\} < 1.0^2, p_e = 0.5$ USING ALFA MODEL

The critical point shifted as the coefficient of variation changed. However, as shown in Fig. 6, for $\{c_a^2, c_s^2\} < 1.0^2, p_e = 0.5$, the positions of the critical points and the values of $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$ were consistent between the approximate formula and extended Alfa model, with a difference of 0.18 [%]. Although the degree of agreement varies depending on the input parameter regions $\{c_a^2, c_s^2\}$, this confirms that the model closely aligns with the results calculated under the single-server conditions in (5) and (6), even when the excluded volume effect is applied to the Alfa model.

TABLE 1. GI/PH/c queuing system with excluded volume effect.

$\rho = 0.4$		2-servers			3-servers			4-servers		
c_a	c_s	formula*	Alfa model	diff%**	formula*	Alfa model	diff%**	Formula*	Alfa model	diff%**
1.0	1.0	3.61090	3.04520	15.67%	3.34180	2.48180	25.74%	3.10970	2.01710	35.14%
0.5	1.0	2.55350	2.41500	5.42%	2.39170	1.89730	20.67%	2.25570	1.49020	33.94%
1.0	0.5	1.63040	1.58580	2.74%	over array			over array		
0.5	0.5	0.73680	1.04280	41.52%	0.67800	0.68311	0.75%	0.62950	0.44723	28.96%
0.3	1.0	2.23570	1.95970	12.34%	2.09410	1.48070	29.29%	1.97530	1.11840	43.38%
1.0	0.3	1.40210	1.29340	7.75%	1.27840	0.89828	29.74%	1.17960	0.62334	47.16%
0.3	0.3	over array			over array			over array		
0.7	1.0	2.95580	3.45220	16.79%	2.76440	2.23690	19.08%	2.60290	1.80050	30.83%
1.0	0.7	2.11880	2.00960	5.16%	1.94450	1.53700	20.96%	1.80400	1.17540	34.84%
0.7	0.7	1.53160	1.74900	14.20%	1.41400	1.29700	8.27%	1.31730	0.96159	27.01%

Formula * is from the Sakasegawa approximations (5) for multiple servers, with the excluded volume effect applied; diff%**: % conversion of the absolute difference between formula* and the Alfa model.

2) CASE: $\left\{c_a^2, c_s^2\right\} \leq 1.2^2, p_e = 0.5$ using Alfa model

For $\left\{c_a^2, c_s^2\right\} \leq 1.2^2$, the gamma distribution has an exceptionally long slope because the mean value is smaller and the coefficient of variation is larger. This is considered to be an issue of computational resources and results in good agreement with the approximation formula that can be obtained if the numerical calculations retain sufficient accuracy. Because an exceptionally large array size is required to accurately represent such a distribution, the limited range of $\left\{c_a^2, c_s^2\right\} \leq 1.2^2$ is presented here for validation. As shown in Fig. 7, the dependence of $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$ on $\left\{c_a^2, c_s^2\right\}$ was qualitatively well reproduced.

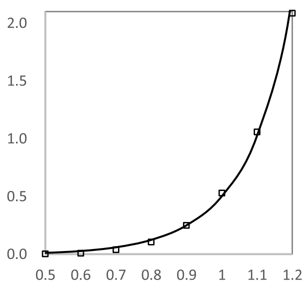

FIGURE 7. When $c_a^2 = \left\{0.5^2 \sim 1.2^2\right\}, c_s^2 = 1.0^2, \rho = 0.25, dt = \{0.1, 1\}$, and$p_e = 0.5$: Comparison of $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$ between (5) and the Alfa model (Vertical axis: $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$, horizontal axis: c_a^2, solid curve: from (5), dots: from Alfa model).

C. VALIDATION WITH GI/PH/C QUEUING SYSTEM WITH EXCLUDED VOLUME EFFECT

Table 1 summarizes the results of the experimental model verification in which the excluded volume effect was adapted to the aggregated Alfa model. Under conditions with small values of c_a^2, c_s^2, and ρ, as the number of servers increases, the array size tends to increase, and the diff% also tends to increase.

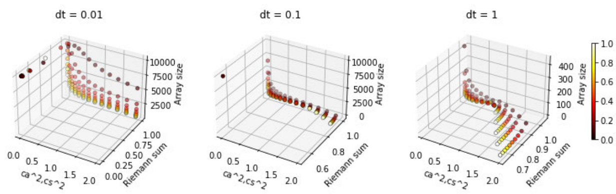

FIGURE 8. Required array sizes dt = {0.01, 0.1, 1} with each $\left\{c_a^2, c_s^2\right\} \leq 2^2$ and $\rho \leq 1$. x-axis: c_a^2, c_s^2, Y-axis: Riemann sum, Z-axis: array size. Left dt = 0.01, center dt = 0.1, and right dt = 1.

D. DISCUSSIONS: LIMITATIONS AND ASSUMPTIONS

As summarized in Table 1, diff%** (% conversion of the absolute difference between (5) and the Alfa model) is strongly influenced by the number of arrays, squared value of the coefficient of variation, and number of servers. In particular, the array size tends to increase under small values of the coefficient of variation and small values of ρ. Furthermore, when the number of servers increases, the array size must be increased proportionally, following the relation (number of servers) $\times$ (array size). The typical cloud computing environment we have contracted with supports of approximately 600 - 700 array sizes. At this time, validating beyond this is impractical.

Figure 8 shows the required array sizes for $\left\{c_a^2, c_s^2\right\} \leq 2$, and $\rho \leq 1$, with dt $= \{0.01, 0.1, 1\}$. The optimal array size is predicted using the array size, where the Riemann sum is equals to one for the gamma distribution. Most cases in which the Riemann sum equals 1 occur with dt $= 0.01$, where array sizes often exceed 700. Therefore, dt $= 0.1$ is selected for practical verification.

VII. RESULTS AND CONCLUSION

In this study, we propose the Alfa model, which integrates the excluded volume effect in the TASEP model into the

TABLE 2. List of state transitions with master equations.

State transitions	Terms of the master equations	State descriptions	The master equations
$P_A(1) \to P_A(1)$	$(1-\lambda)(1-\mu)P_A(1)$	No arrival, No processing	$P_A(1) = (1-\lambda)(1-\mu)P_A(1) + \lambda P_B(0) + (1-\lambda)P_B(1).$
$P_B(0) \to P_A(1)$	$\lambda P_B(0)$	Arriving	
$P_B(1) \to P_A(1)$	$(1-\lambda)P_B(1)$	No arrival	
$P_A(n-1) \to P_A(n)$	$\lambda(1-\mu)P_B(1)$	Arriving, No processing	$P_A(n) = \lambda(1-\mu)P_B(1) + (1-\lambda)(1-\mu)P_A(n) + \lambda P_B(n-1) + (1-\lambda)P_B(n).$ $(n \geq 2)$
$P_A(n) \to P_A(n)$	$(1-\lambda)(1-\mu)P_A(n)$	No arrival, No processing	
$P_B(n-1) \to P_A(n)$	$\lambda P_B(n-1)$	Arriving	
$P_B(n) \to P_A(n)$	$(1-\lambda)P_B(n)$	No arrival	
$P_B(0) \to P_B(0)$	$(1-\lambda)P_B(0)$	No arrival	$P_B(0) = P(0) = (1-\lambda)P_B(0) + (1-\lambda)\mu P_A(1).$
$P_A(1) \to P_B(0)$	$(1-\lambda)\mu P_A(1)$	No arrival, No processing	
$P_A(n) \to P_B(n)$	$\lambda\mu P_A(n)$	Arriving, Processing	$P_B(n) = \lambda\mu P_A(n) + (1-\lambda)\mu P_A(n+1).$ $(n \geq 1)$
$P_A(n+1) \to P_B(n)$	$(1-\lambda)\mu P_A(n+1)$	No arrival, Processing	

numerical calculations for a discrete-time GI/G/1 system based on the MGM. This model was devised by interpreting the delay in the progression of the matrix as probability. This was subsequently incorporated into the numerical calculations for the GI/G/1 queue. In addition, we derive an approximation formula (5) to express the queue length $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$. For validation, we performed numerical calculations for the queue length $L_{q(\mathrm{GI/G/1})\mathrm{ASEP}}$ in the GI/G/1 model with the excluded volume effect for $\{c_a^2, c_s^2\} \leq 1.2^2$. The results of the proposed approximation formula and the extended Alfa model exhibited good agreement within a limited range of $\{c_a^2, c_s^2\}$ values and their combinations. Moreover, the verification of GI/PH/c using the aggregation method with the Alfa model revealed high validity ratings, although within a limited range of $s \leq 4$. Verification across multiple servers is challenging because of insufficient computing resources. However, because the MGM uses the same algorithm, the validity of the calculated results can be verified. The extended Alfa model is a highly versatile numerical method suitable for detailed analyses across a broad range of systems. The proposed method for handling the excluded volume effect is considered effective in introducing this effect into general GI/G/1 systems.

The proposed approximation formula presented herein, along with the use of the Alfa model adapted to the excluded volume effect are expected to be applicable to models that incorporate physical concepts. The approximation formula is based on the lag between dynamic phenomena and generally distributed stochastic events. This implies that it is possible to model mixed physical phenomena. For instance, we consider complex phenomena that combine multiple processes and factors (e.g., multiphase fluid flow, mass transport in cells, diffusion of particles in liquids, and arrival time intervals of entire systems with random inputs from different sources). An approach employing a general distribution should generate realistic predictions. We expect that this will advance the elucidation of physical phenomena that cannot be represented using conventional models.

APPENDIX

A. DETAILS OF THE CALCULATION FOR (1)

From the literature [3], [10], to derive (1), it is necessary to determine the state transitions, as listed in Table 2, and obtain the recurrence equation.

$P_A(1)$: State A is the probability of a stationary state with the receptor occupied by one particle.

$P_A(n)$: State A is the probability of stationary states with the receptor occupied by one particle and n particles in queue. $(n \geq 2)$

$P_B(1)$: State B is the probability of stationary states with a vacant receptor, and a state in which one particle occupies one cell before the receptor.

$P_B(n)$: State B is the probability of a stationary state with a vacant receptor and n particles in queue. $(n \geq 1)$

$P(0)$: Zero-state is the probability of a stationary state with no particles in the queue.

$P_B(0)$: State B is the probability of stationary states with a vacant receptor equivalent to $P_B(0) = P(0)$.

n: Number of particles in the deterministic movement in the queue; two vacant cells in the queue never appear in the stationary state. Thus, 2^n states were defined with n particles in the queue.

The recurrence formula was derived and rearranged to obtain probability distributions $P(0)$, $P_A(n)$, and $P_B(n)$ (See Appendix 5. D.1 in [10] for derivation). $\rho = \lambda/\mu$.

$$P_A(n) = \left\{\frac{1-\mu+\lambda\mu}{(1-\lambda)^2}\rho\right\}^{n-1} \frac{\lambda}{(1-\lambda)\mu} P(0). \ (n \geq 1)$$

$$P_B(n) = \left\{\frac{1-\mu+\lambda\mu}{(1-\lambda)^2}\rho\right\}^{n-1} \frac{\lambda^2}{(1-\lambda)^2\mu} P(0). \ (n \geq 1)$$

$$P(0) = 1 - \frac{\rho}{1-\lambda}.$$

In that study, the probability distribution $P_L(\ell)$ with queue length ℓ was defined as follows:

$$\tilde{\rho} = \lambda/\tilde{\mu},$$
$$\tilde{\mu} = (1-\lambda)\,\mu,$$
$$P_L(\ell) = (1-\tilde{\rho})\,\tilde{\rho}^{\ell}.\ (\ell \geq 0)$$

Referring to [10, formula means L_q (5.66) and (5.20), respectively] and obtain the following:

$$L_s = \sum\nolimits_{l=0}^{\infty} \ell P_L(\ell),$$

and,

$$P_L(\ell) = (1-\tilde{\rho})\,\tilde{\rho}^{\ell}.\ (1 \geq 0)$$

Apply $P_L(\ell)$ to L_q as follows:

$$L_q = \sum_{l=0}^{\infty} (\ell - 1)\, P_L(l) = \sum_{l=0}^{\infty} (\ell - 1)(1-\tilde{\rho})\,\tilde{\rho}^{l}.$$

Apply $\ell = 0\ to\ \infty$, and subtract recurrences.

Then,

$$L_q = (1-\tilde{\rho})\,\tilde{\rho}^2 \frac{d}{d\tilde{\rho}} \sum_{l=0}^{\infty} \tilde{\rho}^{\ell-1} = \frac{\tilde{\rho}^2}{1-\tilde{\rho}}.$$

Check the denominator in L_q. The critical line is $1-\tilde{\rho} = 0$ and then,

$$\tilde{\rho} = \frac{\lambda}{\tilde{\mu}} = \frac{\lambda}{(1-\lambda)\,\mu} = 1.$$

Obtaining (1), where λ_{cr} is the critical value of λ. This satisfies the $\lambda_{cr} > \lambda$ relationship when the stationary state of the queue is under an excluded volume effect with parallel update states. In this regard, this relationship is important.

$$\lambda_{cr} = \frac{\mu}{1+\mu}.$$
$$(\lambda_{cr} < \mu\ for\ 0 < \mu \leq 1)$$

B. MGM [19] AND ALFA MODEL [22]

The Alfa model is a geometric matrix for discrete-time GI/G/1 queuing systems. The state of the system at time n is $\{L_n, K_n, J_n\}$ or components (i, k, j), which are a combination of the number of particles in system L_n at the time remaining until the next particle arrives K_n and the time remaining until the particle currently being processed exits J_n. The set of states considered by the system is as follows:

$$\Delta = \{(i,k,j)\,; i=0,1,2,\cdots, k=1,2,\cdots,n_t, j=1,2,\cdots,n_s\}$$
$$(\text{when } i=0, j=0).$$

The state of the system at time n is defined as the row vector $\boldsymbol{x} = [x_0, x_1, x_2, \cdots], (i = \{0, 1, 2, \cdots\})$. Let the state of the system at time n be the number of particles in system L_n, the time remaining until the next particle arrives K_n, the time remaining until the particle currently being processed leaves J_n, and let $\boldsymbol{x_i}$ be the state vector. The state components (i, k, j) of each particle represent the probability that the particle will take the state defined by the component, where $i = L$ denotes the number of particles in the system, $k = \alpha_k$ denotes the state probability of the remaining time until the particle arrives, and $j = \boldsymbol{\beta}_j$ denotes the remaining processing time. For the transition matrix P_A representing the time evolution of state vector $\boldsymbol{x}_i$, representing the state change of the arrival time of the particle, row vector $\boldsymbol{\alpha} = \left[\alpha_1, \alpha_2, \cdots, \alpha_{n_t}\right]$ representing the arrival time interval distribution representing the remaining time until the particle arrives, and matrix T representing the time-state transition, the arrival can be represented as follows (the process side has the same structure but different parameters):

$$T = \begin{bmatrix} 0_{n_t-1} & 0 \\ I_{n_t-1} & 0^T_{n_t-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$
$$t = 1 - T1,$$
$$P_A = \begin{bmatrix} \mathbf{1} & 0 \\ \boldsymbol{t} & T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The vector representing the state of the system is defined as $\boldsymbol{x}_A^{[0]} = [0\ 0\ 0\ 1\ 0]$, where each component (i, k, j) represents the probability of being in state (i, k, j) and the time evolution of the system is determined as $\boldsymbol{x}_A^{[1]} = \boldsymbol{x}_A^{[0]} P_A$. Vector $\boldsymbol{\pi}$ represents the steady state, satisfies $\boldsymbol{\pi} = \boldsymbol{\pi} P$, and its state is computed following the MGM computation rules. The transition matrix P representing the change in state x of the system is expressed as follows:

$$P = \begin{bmatrix} B & C & 0 & 0 & \cdots \\ E & A_1 & A_0 & 0 & \cdots \\ 0 & A_2 & A_1 & A_0 & \ddots \\ 0 & 0 & A_2 & A_1 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

where,

$$B = T,$$

$$
\begin{aligned}
C &= \beta \otimes (t\alpha)\,,\\
E &= s \otimes T,\\
A_0 &= S \otimes (t\alpha)\,,\\
A_1 &= (s\beta) \otimes (t\alpha) + S \otimes T,\\
A_2 &= (s \otimes \beta) \otimes T.
\end{aligned}
$$

By dividing vector $\boldsymbol{\pi}$ into blocks by the number of particles in the system, $\boldsymbol{\pi} = [\pi_0, \pi_1, \pi_2, \pi_3, \cdots]$, and introducing it into the rate matrix R (see below), we obtain the following:

$$
R = A_0 + RA_1 + R^2 A_2 \,(\forall \boldsymbol{\pi}_i)\,,
$$

$$
R = \begin{bmatrix} R_1 & R_2 & \cdots & R_{n_t} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.
$$

Given the following relation:

$$
\begin{aligned}
\boldsymbol{\pi}_i &= \boldsymbol{\pi}_{i-1} A_0 + \boldsymbol{\pi}_i A_1 + \boldsymbol{\pi}_{i+1} A_2\\
&\qquad (i \geq 2, i = \{0, 1 \cdots\})\\
\boldsymbol{\pi}_{i+1} &= \boldsymbol{\pi}_i R,\\
&\qquad (i \geq 1, R = constant)
\end{aligned}
$$

We can write as follows:

$$
\boldsymbol{\pi}_{i-1} R = \boldsymbol{\pi}_{i-1} A_0 + \boldsymbol{\pi}_{i-1} R A_1 + \boldsymbol{\pi}_{i-1} R^2 A_2.
$$

Here, by computing matrix R, we can write as follows:

$$
\begin{aligned}
R_1 &= a_1 S + a_1 R_1 (s\beta) + R_2 S + R_1 R_2 (s\beta)\,,\\
&\cdots,\\
R_{n_t} &= a_{n_t} S + a_{n_t} R_1 (s\beta)\,.
\end{aligned}
$$

Acting on vector $\boldsymbol{\pi} = [\pi_0, \pi_1, \pi_2, \pi_3, \cdots]$ by matrix R, we obtain the following:

$$
\begin{aligned}
\boldsymbol{\pi}_2 &= \boldsymbol{\pi}_1 R,\\
\boldsymbol{\pi}_0 &= \boldsymbol{\pi}_0 B + \boldsymbol{\pi}_1 E,\\
\boldsymbol{\pi}_1 &= \boldsymbol{\pi}_0 C + \boldsymbol{\pi}_1 (A_1 + RA_2)\,,\\
[\boldsymbol{\pi}_0 \boldsymbol{\pi}_1] &= [\boldsymbol{\pi}_0 \boldsymbol{\pi}_1] \begin{bmatrix} B & C \\ E & A_1 + RA_2 \end{bmatrix},\\
&\qquad (standardized\ condition{:}\boldsymbol{\pi} 1 = 1)
\end{aligned}
$$

Then,

$$
\begin{aligned}
1 &= \boldsymbol{\pi} 1\\
&= \sum\nolimits_{i=0}^{\infty} \boldsymbol{\pi}_i 1\\
&= \boldsymbol{\pi}_0 1 + \boldsymbol{\pi}_1 (I - R)^{-1} 1,\\
\boldsymbol{\pi}_{i+1} &= \boldsymbol{\pi}_i R.\\
&\qquad (t \geq 1, R = constant)
\end{aligned}
$$

The obtained vector $\boldsymbol{\pi} = [\pi_0, \pi_1, \pi_2, \pi_3, \cdots]$ is adapted to the following: $\mu_L = \sum_{k=0}^{\infty} k \pi_k 1$

$$
= \pi_1 (1 - R)^{-2} 1.
$$

The average queue length is μ_L [21].

C. AGGREGATION METHOD IN MGM [22]

The Alfa model [22] defines the time interval for processing GI/G/s queuing systems by superimposing all independent distributions of multiple servers in the processing system of the model using the aggregation method (hereinafter it is described as the aggregated Alfa model). This model can be interpreted as a hyper-Erlang distribution if the shape parameter α of the entire processing distribution superimposed by the aggregation method is an integer value ($\alpha \in \mathbb{Z}$), and as a hyper-exponential distribution if $\alpha = 1$. This distribution is known to be a distribution that can comprehensively reproduce the dynamics of the superimposed gamma distribution (hereinafter referred to as the hyper-Gamma distribution) [22]. Hereinafter, the probability density function (PDF) with shape parameter α and parameter λ is denoted as g (x;α, λ) and can be described as a distribution in which multiple PDFs are superimposed with the weight of probability p_i of taking each distribution, where i denotes the i-th receptor in the aggregation model ($i \in \mathbb{Z}_+$).

$$
\begin{aligned}
f_g(x) &= \sum_{i=1}^{n} p_i g(x; \alpha_i, \lambda_i)\,,\\
g(x; \alpha, \lambda) &= \frac{\lambda^{\alpha} x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x},\\
\Gamma(\alpha) &= \int_0^{\infty} e^{-x} x^{\alpha-1} dx,\\
\Gamma(n) &= (n-1)!, \,(\alpha = n \in Z_+)\,.
\end{aligned}
$$

The entire system can be interpreted as a model that adopts a time interval according to a single distribution, by aggregating the distributions of each server. In a previous study [23], the coefficient of variation of the processing time of a G/G/1 queue was calculated as follows: $c_s^2 = \sigma_s^2 \mu^2$. σ_s denotes the variance in the distribution of the processing time intervals. This relationship was adapted for superposition of the processing-side variation coefficient values of the aggregated Alfa model. The calculation system for queue length $L_{q(\mathrm{GI/PH/c})}$ in the aggregation model is as follows:

$$
\begin{aligned}
L_{q(GI/PH/c)} &= \sum_{k=1}^{\infty} (k-1)\, \boldsymbol{\pi}_k 1\\
&= 0\boldsymbol{\pi}_1 1 + 1\boldsymbol{\pi}_2 1 + 2\boldsymbol{\pi}_3 1 + \cdots\\
&= \boldsymbol{\pi}_1 R \left(I + 2R + 3R^3 + \cdots\right) \mathbf{1}\\
&= \boldsymbol{\pi}_1 R (I - R)^{-2} \mathbf{1},
\end{aligned}
$$

and

$$
\begin{aligned}
L_{q(GI/PH/c)} &= \sum\nolimits_{k=s}^{\infty} (k - s\rho)\, \boldsymbol{\pi}_k 1\\
&= \boldsymbol{\pi}_1 R^{s\rho} (I - R)^{-2} \mathbf{1}.
\end{aligned}
$$

The $L_{q(\mathrm{GI/PH/c})}$ calculation system in the aggregated Alfa model considers the dependence of traffic intensity in the queuing system and estimates the average number of particles processed at the receptors by $s\rho$ [15]. The term $R^{s\rho}$

TABLE 3. Performance evaluations of (2) and (5) (case: $\left\{c_a^2, c_s^2\right\} \leq 1.0^2, p_e = \{0.5, 1.0\}, dt = 0.1$) [22], [27].

input parameters			$p_e = 0.5$			$p_e = 1.0$			$dt = 0.1$	
ρ	C_a	C_s	eq. (5)	Alfa model	diff%	eq. (2)	Alfa model	diff%	totals array size*	Average Riemann sum
0.35	1.00	1.00	1.83841	1.71710	6.60%	0.20517	0.18537	9.65%	394	0.99995
0.40	1.00	1.00	3.86153	3.62150	6.22%	0.28963	0.26166	9.66%	344	0.01000
0.45	1.00	1.00	11.37710	10.65750	6.32%	0.39898	0.36040	9.67%	306	0.99992
0.48	1.00	1.00	65.21770	60.85680	6.69%	0.47883	0.43281	9.61%	286	0.99990
0.35	0.30	1.00	1.08056	1.05960	1.94%	over array			891	0.99998
0.40	0.30	1.00	2.30472	2.40090	4.17%	0.16611	0.11684	29.66%	779	0.99997
0.45	0.30	1.00	7.94780	8.69120	9.35%	0.23119	0.17594	23.90%	692	0.99996
0.48	0.30	1.00	unstable rho			0.23119	0.17594	23.90%	649	0.99995
0.35	1.00	0.30	over array			over array			891	0.99998
0.40	1.00	0.30	1.56283	1.85780	18.87%	0.14041	0.15315	9.07%	779	0.99997
0.45	1.00	0.30	3.53147	4.20450	19.06%	0.19212	0.20943	9.01%	692	0.99996
0.48	1.00	0.30	7.78561	9.24010	18.68%	0.22991	0.25048	8.94%	649	0.99995
0.35	0.30	0.30	over array			over array			1388	1.00000
0.40	0.30	0.30							1214	1.00000
0.45	0.30	0.30							1078	1.00000
0.48	0.30	0.30							1012	1.00000
0.35	0.50	1.00	1.23684	1.29630	4.81%	0.13359	0.10913	18.31%	570	0.99998
0.40	0.50	1.00	2.75291	3.05840	11.10%	0.19022	0.16389	13.84%	498	0.99997
0.45	0.50	1.00	9.00907	10.08070	11.90%	0.26462	0.23819	9.99%	443	0.99996
0.48	0.50	1.00	unstable rho			0.32010	0.29479	7.91%	415	0.99995
0.35	1.00	0.50	0.90807	1.06230	16.98%	0.11398	0.12704	11.46%	570	0.99998
0.40	1.00	0.50	1.81704	2.17190	19.53%	0.15995	0.17827	11.45%	498	0.99997
0.45	1.00	0.50	4.16734	4.98260	19.56%	0.21880	0.24385	11.45%	443	0.99996
0.48	1.00	0.50	9.32014	11.13400	19.46%	0.26172	0.29164	11.43%	415	0.99995
0.35	0.50	0.50	over array			over array			746	1.00000
0.40	0.50	0.50	0.81362	1.58780	95.15%	0.09274	0.13714	47.88%	652	1.00000
0.45	0.50	0.50	2.08577	4.22350	102.49%	0.09274	0.13714	47.88%	580	1.00000
0.48	0.50	0.50	6.16812	12.73300	106.43%	0.11160	0.16978	52.13%	544	1.00000
0.35	0.70	1.00	1.45230	1.50580	3.68%	0.15791	0.14296	9.47%	461	0.99998
0.40	0.70	1.00	3.01517	3.16920	5.11%	0.22436	0.20840	7.11%	403	1.00002
0.45	0.70	1.00	9.65001	10.29880	6.72%	0.31132	0.29538	5.12%	358	1.00001
0.48	0.70	1.00	208.42297	215.94920	3.61%	0.37593	0.36069	4.05%	335	1.00000
0.35	1.00	0.70	1.20782	1.34710	11.53%	0.14032	0.14809	5.54%	461	0.99998
0.40	1.00	0.70	2.32488	2.59160	11.47%	0.19705	0.20804	5.58%	403	1.00002
0.45	1.00	0.70	5.61628	6.25700	11.41%	0.27029	0.28522	5.53%	358	1.00001
0.48	1.00	0.70	14.12738	15.73030	11.35%	0.32370	0.34155	5.51%	335	1.00010
0.35	0.70	0.70	0.83973	1.14510	36.36%	0.09531	0.10940	14.78%	528	1.00000
0.40	0.70	0.70	1.66902	2.32260	39.16%	0.13480	0.15913	18.05%	462	1.00010
0.45	0.70	0.70	4.41974	6.25510	41.53%	0.18604	0.22481	20.84%	410	1.00010
0.48	0.70	0.70	14.64317	20.90010	42.73%	0.22380	0.27381	22.35%	384	1.00000

diff%**: % conversion of the absolute difference between (5), Alfa model; n/a*: unstable stability, and total array size*: total number of required arrays of inflow and outflow probability distributions.

in the same equation yields complex results. As a countermeasure, for a server number s, the following equation $L_{q(m)} = \sum_{k=s}^{\infty}(k-m)\boldsymbol{\pi}_k 1 = \boldsymbol{\pi}_1 R^m (I-R)^{-2}\mathbf{1}$ is solved for each integer value m in the neighborhood, and the number of particles present $L_{q(\mathrm{GI/PH/c})}$ is obtained by interpolating both.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Daichi Yanagisawa, Associate Professor, from The University of Tokyo, Prof. Stefania Bandini from the University of Milano-Bicocca, and The University of Tokyo, and Toshio Akimitsu, Associate Professor, from The Open University of Japan for their advice on this study. They would also like to thank the authors, Prof. Attahiru Sule Alfa (University of Manitoba and University of Pretoria) and Prof. Wei Wayne Li (Texas Southern University) for introducing helpful references [19], [20], and [21]. They also like to thank Dr. Hironori Shimoyama and Editage for English language editing, the Microsoft Azure support team for the cloud computing system, and my family.

REFERENCES

[1] A. B. Kolomeisky, G. M. Schütz, E. B. Kolomeisky, and J. P. Straley, "Phase diagram of one-dimensional driven lattice gases with open boundaries," *J. Phys. A, Math. Gen.*, vol. 31, no. 33, pp. 6911–6919, Aug. 1998, doi: 10.1088/0305-4470/31/33/003.

[2] A. Schadschneider, D. Chowdhury, and K. Nishinari, *Stochastic Transport in Complex Systems: From Molecules To Vehicles*. Amsterdam, The Netherlands: Elsevier, 2010.
[3] D. Yanagisawa, A. Tomoeda, R. Jiang, and K. Nishinari, "Excluded volume effect in queueing theory," *JSIAM Lett.*, vol. 2, pp. 61–64, Jun. 2010, doi: 10.14495/jsiaml.2.61.
[4] C. T. MacDonald and J. H. Gibbs, "Concerning the kinetics of polypeptide synthesis on polyribosomes," *Biopolymers*, vol. 7, no. 5, pp. 707–725, May 1969, doi: 10.1002/bip.1969.360070508.
[5] C. Arita and D. Yanagisawa, "Exclusive queueing process with discrete time," *J. Stat. Phys.*, vol. 141, no. 5, pp. 829–847, Dec. 2010, doi: 10.1007/s10955-010-0075-9.
[6] C. Arita, "Queueing process with excluded-volume effect," *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.*, vol. 80, no. 5, Nov. 2009, Art. no. 051119, doi: 10.1103/physreve.80.051119.
[7] C. Arita and A. Schadschneider, "Dynamical analysis of the exclusive queueing process," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 83, no. 5, May 2011, Art. no. 051128, doi: 10.1103/physreve.83.051128.
[8] C. Arita and A. Schadschneider, "Exclusive queueing processes and their application to traffic systems," *Math. Models Methods Appl. Sci.*, vol. 25, no. 3, pp. 401–422, Mar. 2015, doi: 10.1142/s0218202515400011.
[9] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, "Exact solution of a 1D asymmetric exclusion model using a matrix formulation," *J. Phys. A, Math. Gen.*, vol. 26, no. 7, pp. 1493–1517, Apr. 1993, doi: 10.1088/0305-4470/26/7/011.
[10] D. Yanagisawa, "Modeling and analysis on bottlenecks in airplanes and airports towards achievement of smoother pedestrian flow," (in Japanese), Ph.D. dissertation, Dept. Civil Eng., Univ. Tokyo, Tokyo, Japan, 2010. [Online]. Available: http://gakui.dl.itc.u-tokyo.ac.jp/cgi-bin/gazo.cgi?no=126400
[11] T. M. Liggett, *StochastIc Interacting Systems: Contact, Voter and Exclusion Processes*. Berlin, Germany: Springer, 2013.
[12] G. Schütz and E. Domany, "Phase transitions in an exactly soluble one-dimensional exclusion process," *J. Stat. Phys.*, vol. 72, no. 1, pp. 277–296, Jul. 1993, doi: 10.1007/bf01048050.
[13] T. Hoshiyama, "Performance evaluation algorithm for GI/G/m queues," in *Proc. Teletraffic Era Beyond-5G AI (ITC)*, Osaka, Japan, 2020, pp. 1–2.
[14] W. J. Hopp and M. L. Spearman, *Factory Physics*. Long Grove, IL, USA: Waveland Press, 2011.
[15] W. Whitt, "The queueing network analyzer," *Bell Syst. Tech. J.*, vol. 62, no. 9, pp. 2779–2815, Nov. 1983, doi: 10.1002/j.1538-7305.1983.tb03204.x.
[16] H. Sakasegawa, "An approximation formula $L_q \simeq \alpha \cdot \rho^{\beta}/(1\text{-}\rho)$," *Ann. Inst. Stat. Math.*, vol. 29, no. 1, pp. 67–75, Dec. 1977, doi: 10.1007/bf02532775.
[17] E. Page, *Queueing Theory in OR*. London, U.K.: Butterworths, 1972.
[18] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1981.
[19] A. S. Alfa, *Queueing Theory for Telecommunications: Discrete Time Modeling of a Single Node System*. Berlin, Germany: Springer, 2010, doi: 10.1007/978-1-4419-7314-6.
[20] A. S. Alfa and W. Li, "Matrix-geometric analysis of the discrete time Gi/G/1 system," *Stochastic Models*, vol. 17, no. 4, pp. 541–554, 2001, doi: 10.1081/stm-120001222.
[21] A. S. Alfa, "Combined elapsed time and matrix-analytic method for the discrete time GI/G/1 and GI^X/G/1 systems," *Queueing Syst.*, vol. 45, no. 1, pp. 5–25, 2003, doi: 10.1023/A:1025653200299.
[22] T. Hoshiyama, "MGM for aggregated general distribution queuing system in discrete time," in *Proc. IEEE 13th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Mar. 2023, pp. 5–8.
[23] C. C. Roberto, "Approximation for single-channel multi-server queues and queuing networks with generally distributed inter-arrival and service times," Doctoral dissertation, Dept. Eng. Manag. Syst. Eng., Missouri Univ. Sci. Technol., Rolla, MO, USA, 2016. [Online]. Available: https://scholarsmine.mst.edu/doctoral_dissertations/2644/
[24] W. Whitt, "Approximations for the GI/G/m queue," *Prod. Oper. Manage.*, vol. 2, no. 2, pp. 114–161, Jun. 1993.
[25] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," *J. de Phys. I*, vol. 2, no. 12, pp. 2221–2229, Dec. 1992, doi: 10.1051/jp1:1992277.
[26] N. Rajewsky, L. Santen, A. Schadschneider, and M. Schreckenberg, "The asymmetric exclusion process: Comparison of update procedures," *J. Stat. Phys.*, vol. 92, no. 1, pp. 151–194, 1998, doi: 10.1023/a:1023047703307.
[27] T. Hoshiyama, "Cross analysis by matrix geometric method for general distribution probability in discrete time," in *Proc. IEEE World AI IoT Congr. (AIIoT)*, Jun. 2023, pp. 256–258, doi: 10.1109/AIIoT58121.2023.10174297.

TAKAKO HOSHIYAMA (Graduate Student Member, IEEE) received the B.E. degree in industrial chemical process engineering from Kanagawa Institute of Technology, Japan, in 1994, the M.B.A. (Master of Business Administration) degree from the Aoyama Business School, Japan, in 2007, the M.S. degree in natural and environmental sciences from The Open University of Japan, in 2016, and the P.Eng. degree in industrial engineering, Japan, in 2020. She is currently pursuing the Ph.D. degree in advanced science and technology with The University of Tokyo, Japan.

KATSUHIRO NISHINARI has been a Professor with The University of Tokyo, since 2009, and a member of the Physical Society of Japan and Applied Mathematical Society of Japan. He works on interdisciplinary topics in applied mathematics, fluid dynamics, and statistical physics. He has published several research papers in leading international journals and several books concerning traffic jams and applied mathematics. His current research interests include "Jamology," an interdisciplinary area of research created by himself approximately 20 years ago on transportation and jamming phenomena including vehicular traffic, pedestrian motion, queue networks, and supply chain. He has won awards such as the 23rd Scientific Publication Award in Japan by the book "Jamology" (in Japanese) and the IgNovel Prize, in 2021.