Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds.

REGENERATIVE SIMULATION FOR MULTICLASS OPEN QUEUEING NETWORKS

Sarat Babu Moka
Sandeep Juneja

School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai-400005, INDIA

ABSTRACT

Conceptually, under restrictions, multiclass open queueing networks are positive Harris recurrent Markov processes, making them amenable to regenerative simulation for estimating the steady-state performance measures. However, regenerations in such networks are difficult to identify when the interarrival times are generally distributed. We assume that the interarrival times have exponential or heavier tails and show that such distributions can be decomposed into mixture of sums of independent random variables such that at least one of the components is exponentially distributed. This allows an implementable regenerative simulation for these networks. We show that the regenerative mean and standard deviation estimators are consistent and satisfy a joint central limit theorem. We also show that amongst all such interarrival decompositions, the one with largest mean exponential component minimizes the asymptotic variance of the standard deviation estimator. We also propose a regenerative simulation method that is applicable even when the interarrival times have superexponential tails.

1 INTRODUCTION

A regenerative process is a stochastic process with a sequence of random time instants (known as regeneration times) such that at these instants the process probabilistically regenerates itself (Smith 1955; Asmussen 2003). Regenerative simulation exploits this structure by generating independent and identically distributed ($i.i.d.$) cycles via simulation and extracting consistent estimators of the steady-state performance measures from them (see, e.g, Glynn and Iglehart (1987); Henderson and Glynn (2001); Glynn (2006)). Typically, regenerative simulation is considered applicable to queueing networks when the interarrival times are exponentially distributed, where for example, instants of arrivals to an empty network denote a sequence of regeneration times. To see some possible applications of the regenerative simulation to queueing networks, refer, e.g, Iglehart and Shedler (1979; 1980; 1983). Dai (1995) establishes that under stability of the fluid limit model of a multiclass network, the associated Markov process is positive Harris recurrent that guarantees the existence of regenerations in the process (Athreya and Ney 1978; Nummelin 1978). However, the identification of the regeneration instants involves the explicit knowledge of the transition kernel that is typically difficult to compute (see, e.g, Henderson and Glynn (1999)). A-priori it is not clear whether implementable regeneration schemes can be identified for a queueing network when the interarrival times are generally distributed. In this paper we construct an implementable regenerative simulation method to estimate the steady-state performance measures of multiclass open queueing networks (Dai and Meyn 1995; Dai 1995) focusing primarily in the case where the interarrival times to the system are generally distributed and have exponential or heavier tails.

We develop two methods to implement regenerative simulation for these networks. The first method is applicable when the interarrival times are generally distributed with exponential or heavier tails. This implementation is possible based on the observation that random variables with exponential or heavier tails can be re-expressed as a mixture of sums of independent random variables where at least one of the components is exponentially distributed. Networks where the interarrival times have exponential or heavier tails, e.g., Gamma or Hyper-exponentially distributions, are common in practice (see, e.g, Li (1997); Miller and Bhat (1997)). The second method can be applicable even when the interarrival times have superexponential distributions. Finite moment conditions on regenerative cycles are important as they allow construction of asymptotically valid confidence intervals (see, e.g, Glynn and Iglehart (1987)). Our another contribution is to show that, under mild stability conditions, finite moments of the interarrival and service times are sufficient to guarantee the required finite moment conditions on the proposed regenerative cycles.

As is well known, classical regenerative processes are stochastic processes that can be viewed as concatenation of $i.i.d.$ cycles. These have been extended to general regenerative processes that allow some dependence between adjacent cycles (see, e.g, Asmussen (2003)). Despite no proper rule to identify the regenerative structure that minimizes the *asymptotic variance of the standard deviation estimator (AVSDE)*, Andradóttir, Calvin, and Glynn (1995) show that when one regenerative structure is a subsequence of another, the AVSDE associated with the original sequence is smaller than that associated with the subsequence. We extend this result and show that, in our framework, the selection of the interarrival time decomposition with largest mean exponential component results in the minimum AVSDE.

The remaining paper is organized as follows. In Section 3, we review both the classical and general regenerative simulation, and the associated central limit theorems. In Section 4, we show that under mild conditions, a random variable with an exponential or heavier tail distribution can be re-expressed as a mixture of sums of independent random variables where one of the constituent random variables has an exponential distribution. Formal construction of multiclass network is presented in Section 5. Section 6 presents a regenerative simulation methodology for these networks and establishes the finiteness of moments of regeneration intervals under appropriate assumptions on network primitives. In Section 7, we establish that if one regenerative structure is a subsequence of another then the AVSDE associated with the former regenerative structure is at least as large as that associated with the latter, and present an important application of this result. Section 8 illustrates the proposed simulation method using a simple numerical example. In Section 9, we propose a different method of regenerative simulation which can be applicable even when the interarrival times have superexponential tail distributions. We refer the reader to (Moka and Juneja 2013) for proofs of the results presented in this paper.

2 NOTATION AND TERMINOLOGY

We first introduce notation that will be used throughout the paper. Assume that all the random variables and stochastic processes are defined on a common probability space $(\Omega, \mathscr{F}, P)$. For any metric space $\mathscr{S}$, $\mathscr{B}_{\mathscr{S}}$ denotes the Borel σ-algebra on it. The notation $X \sim F$ is used to denote that the distribution of a random variable X is F. $Exp(\lambda)$ represents an exponential distribution with rate $\lambda > 0$ and $Uniform(a,b)$ represents a uniform distribution on interval $[a,b]$. We write $X_1 \stackrel{d}{=} X_2$ to denote the equivalence of distributions of random variables X_1 and X_2. A subprobability measure ν is a *component* of the distribution of a $\mathscr{S}$-valued random variable X if $P(X \in A) \geq \nu(A)$ for every Borel set $A \in \mathscr{B}_{\mathscr{S}}$ (refer to Thorisson (2000)). The indicator function is denoted by $\mathbb{I}(\cdot)$, which is 1 if the argument is true and 0 otherwise. We say that a probability distribution is *lattice* if it is concentrated on a set of points of the form $a + nh$, where $h > 0$, a is a real value and $n = 0, \pm 1, \pm 2, \ldots$. Any non-zero σ-finite measure π is an *invariant* measure of the process $X = \{X(t) : t \geq 0\}$ if $\pi(B) = \int_{\mathscr{X}} P(X(t) \in B / X(0) = y)\, \pi(dy)$, for all $B \in \mathscr{B}_{\mathscr{X}}, t \geq 0$, where $\mathscr{X}$ is the state space of X. If every invariant measure is a positive scalar multiple of π, then it is well known that π is the *unique* invariant probability measure of the process X. We assume that all the processes considered in this paper are càdlàg (right continuous paths with left limits).

3 REGENERATIVE PROCESSES

As is well known, a sequence of random variables $0 = T_{-1} \leq T_0 < T_1 < T_2 < \cdots$ is called a *renewal process* if the sequence of intervals $\{T_n - T_{n-1}; n \geq 1\}$ is an i.i.d. sequence and independent of T_0. The following definition of regenerative process is based on Asmussen (2003) (also refer to Thorisson (1983) where it is known as wide-sense regenerative process).

Definition 3.1 *A stochastic process $Y = \{Y(t) : t \geq 0\}$ is called* regenerative *if there exists a renewal process $0 = T_{-1} \leq T_0 < T_1 < \cdots$ such that (i) $\{Y(T_n + s) : s \geq 0\}$ is independent of $\{T_0, \ldots, T_n\}$ and (ii) $\{Y(T_n + s) : s \geq 0\}$ is stochastically equivalent to $\{Y(T_0 + s) : s \geq 0\}$ for $n \geq 0$.*

The sequence $T_0, T_1, \ldots$ is referred to as a sequence of *regeneration times*. In addition, if the regeneration cycles, $\{\{Y(s) : T_{n-1} \leq s < T_n\}, n \geq 0\}$, are independent then the process is known as *classically* regenerative. We refer to Y as a non-delayed regenerative process if $T_0 = 0$. Conceptually, one may think of a non-delayed classically regenerative process as a concatenation of $i.i.d.$ cycles.

The following theorem is important to our analysis. Refer, e.g., to Theorem 1.2 in Chapter VI of Asmussen (2003) for proof. Let $\mathscr{Y}$ be the state space of the process Y, and for any distribution ν on $(\mathscr{Y}, \mathscr{B}_{\mathscr{Y}})$, let $\mathbb{E}_\nu(\cdot) := \int_{\mathscr{Y}} \mathbb{E}_y(\cdot)\nu(dy)$, where E_y is the expectation operator associated with the probability measure P_y that satisfies $P_y(Y(0) = y) \equiv 1$.

Theorem 3.1 *Suppose that Y is a non-delayed regenerative process with regeneration times $0 = T_0 < T_1 < \cdots$ and the distribution of the first cycle length T_1 is non-lattice with finite mean. Then, the steady-state distribution π exists and for any non-negative real valued function h, $\mathbb{E}_\pi[h(Y(t))] = \frac{1}{\mathbb{E}_\varphi[T_1]}\mathbb{E}_\varphi\left[\int_0^{T_1} h(Y(s))\,ds\right]$, where φ is the distribution of the initial state $Y(0)$.*

From Theorem 3.1, it is not hard to see that associated with every regenerative process there exists a classically regenerative process with the same steady-state performance measures. Suppose that $X = \{X(t) : t \geq 0\}$ is such a classically regenerative process associated with the regenerative process Y. Often a steady-state performance measure of interest has the form $\bar{r} := \int_{\mathscr{Y}} h(y)\pi(dy)$, where h is a non-negative real valued function defined on $\mathscr{Y}$.

For regenerative simulation to estimate $\bar{r}$, define $\beta(t) := \frac{\sum_{i=1}^{N(t)} R_i}{\sum_{i=1}^{N(t)} \tau_i}$ and $s(t) := \sqrt{\frac{\sum_{i=1}^{N(t)}(R_i - \beta(t)\tau_i)^2}{\sum_{i=1}^{N(t)} \tau_i}}$, for $t \geq 0$, where the counting process $N(t) := \max\{n \geq 0 : S_n \leq t\}$ which counts the number of regenerations that have occurred till time t, and for each $i \geq 1$, $R_i := \int_{S_{i-1}}^{S_i} h(X(s))\,ds$ and $\tau_i := S_i - S_{i-1}$. Set $W_i = R_i - \bar{r}\tau_i$ for each $i \geq 1$. Refer to Glynn and Iglehart (1987) for a proof of Theorem 3.2.

Theorem 3.2 *Let X be a classically regenerative process with regeneration times $0 = S_0 < S_1 < \cdots$ and invariant probability measure π. Set $\varphi(dy) := P(X(0) \in dy)$ and assume that $h \geq 0$.*

(i) *If $\mathbb{E}_\varphi\tau_1 < \infty$, then $\bar{r} = \frac{\mathbb{E}_\varphi R_1}{\mathbb{E}_\varphi \tau_1}$ and $\beta(t) \longrightarrow \bar{r}$, a.s. as $t \to \infty$*

(ii) *If $\mathbb{E}_\varphi\left[R_1^2 + \tau_1^2\right] < \infty$, then as $t \to \infty$, $s(t) \longrightarrow \sigma$, a.s. and $t^{\frac{1}{2}}\dfrac{(\beta(t) - \bar{r})}{s(t)} \Rightarrow \mathscr{N}(0,1)$,*

(iii) *If $\mathbb{E}_\varphi\left[R_1^4 + \tau_1^4\right] < \infty$ and $\sigma > 0$, then $t^{\frac{1}{2}}(\beta(t) - \bar{r}, s(t) - \sigma) \Rightarrow \mathscr{N}(\vec{0}, \mathscr{K})$, as $t \to \infty$,*

where $\mathscr{K} = \frac{1}{\mathbb{E}_\varphi\tau_1}\begin{bmatrix} \mathbb{E}_\varphi\left(W_1^2\right) & \frac{\mathbb{E}_\varphi[(A_1 - bW_1)W_1]}{2\sigma} \\ \frac{\mathbb{E}_\varphi[(A_1 - bW_1)W_1]}{2\sigma} & \frac{\mathbb{E}_\varphi\left[(A_1 - bW_1)^2\right]}{4\sigma^2} \end{bmatrix}$, $A_i = W_i^2 - \sigma^2\tau_i$, $b = 2\mathbb{E}_\varphi(W_1\tau_1)/\mathbb{E}_\varphi\tau_1$ *and* $\mathscr{N}(\vec{0}, \mathscr{K})$ *represents a multivariate normal random variable with the covariance matrix $\mathscr{K}$. The constant σ^2 is known as the time-average variance constant (TAVC) of $h(X(\cdot))$.*

In the above theorem, (i) constitutes the strong law of large numbers (*SLLN*) for point estimator $\beta(t)$. Part (ii) constitutes the *SLLN* for the standard deviation estimator $S(t)$, and the CLT for $\beta(t)$. From (i) and (ii), the asymptotic $100(1-\delta)\%$ confidence interval is $\left[\beta(t) - \frac{zs(t)}{\sqrt{t}},\ \beta(t) + \frac{zs(t)}{\sqrt{t}}\right]$, where z solves the equation $P(-z \leq \mathcal{N}(0,1) \leq z) = 1-\delta$. Finally, part (*iii*) constitutes the joint *CLT* for $(\beta(t), s(t))$. Now it is clear that to construct valid confidence intervals, one sufficient condition is to ensure $\mathbb{E}_{\varphi}\left[R_1^4 + \tau_1^4\right] < \infty$. To establish these moments, in sections 6.2 and 6.3, we study sufficient moment conditions on interarrival and service times in the context of multiclass networks.

4 EXTRACTING EXPONENTIAL COMPONENT IN A DISTRIBUTION

In this section, we show that under mild conditions, any random variable with an exponential or heavier tail distribution can be re-expressed as a mixture of sums of independent random variables such that one of the constituent random variables has an exponential distribution. We then observe this decomposition for well-known distributions such as Pareto, Weibull, Gamma, etc.

Let f be the probability density function of a real valued random variable. We say that $f \in \mathscr{H}$, if there exists an $a \in [-\infty, \infty)$ such that $f(x) = 0$, for all $x < a$ (if $a > -\infty$), f is differentiable on (a, ∞) and $\lambda_f := \sup_{y \in (a,\infty)} \left(-\frac{f'(y)}{f(y)}\right) \in (0, \infty)$. Examples of densities in $\mathscr{H}$ are discussed later.

For $f \in \mathscr{H}$, define $G_\lambda^f(x) := F(x) + \frac{f(x)}{\lambda},\ \ x \in \mathbb{R}$, where the distribution function $F(x) = \int_{-\infty}^{x} f(y)dy$, and $\lambda > 0$. Theorem 4.1 is one of our key results and the proof primarily depends on the fact that any characteristic function uniquely identifies the associated probability distribution.

Theorem 4.1 *Suppose that ξ is a random variable with density $f \in \mathscr{H}$. If $\lambda \geq \lambda_f$, then G_λ^f is a probability distribution function and*

$$\xi \stackrel{d}{=} E + Z, \tag{4.1}$$

where $E \sim Exp(\lambda)$, $Z \sim G_\lambda^f$ and they are mutually independent.

Theorem 4.1 ensures that if $\lambda_f < \infty$ holds then ξ must have an exponential or heavier tail distribution. Now we consider some practically important classes of distributions with exponential or heavier tails.

Example 4.1 (**Lognormal**) Suppose $\xi = \exp\left[\mathcal{N}\left(\mu, \sigma^2\right)\right]$, that is, ξ has a Lognormal distribution, where $\mathcal{N}\left(\mu, \sigma^2\right)$ represents a normal random variable with mean μ and variance $\sigma^2 > 0$. Then, its pdf $g(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right],\ x > 0$ and $\lambda_g = \sup_{x>0}\left(-\frac{g'(y)}{g(y)}\right) = \frac{\exp\left(\sigma^2 - (\mu+1)\right)}{\sigma^2}$; and it is achieved at $x = \exp\left(\mu + 1 - \sigma^2\right)$. Thus, the mean of the maximum exponential component that can be extracted in this manner equals $\sigma^2 \exp((\mu+1) - \sigma^2)$. As is well known, $\mathbb{E}[\xi] = \exp(\mu + \sigma^2/2)$, so that the ratio of the former to the latter equals $\sigma^2 \exp(1 - 3/2\sigma^2)$. Interestingly, this is independent of μ and is maximized at $2/3$ when $\sigma^2 = 2/3$.

Example 4.2 (**Gamma**) Suppose ξ is $Gamma(\alpha, \gamma)$ distributed with pdf $g(x) = \frac{\gamma^\alpha x^{\alpha-1}\exp(-\gamma x)}{\Gamma(\alpha)}$ for $x \geq 0$, $g(x) = 0$ otherwise, where $\gamma > 0$ and $\alpha \geq 1$. Then, it is easily seen that $\lambda_g = \sup\left(-\frac{g'(y)}{g(y)}\right) = \gamma$ and hence, we achieve the desired decomposition.

Example 4.3 (**Log-Convex densities, Pareto**) Let ξ be a random variable with log-convex density g that has support $[a, \infty)$ and is differentiable on (a, ∞). Since $-\log g(x)$ is concave function of x, $\frac{-g'(x)}{g(x)} = \frac{d}{dx}(-\log g(x))$ is a monotonically non-increasing function of x (because $\frac{d^2}{dx^2}(-\log g(x)) \leq 0$). Furthermore, $\lambda_g = \frac{-g'(a)}{g(a)}$ and when $\lambda_g \in (0, \infty)$, we achieve the desired decomposition.

For instance, suppose that ξ has Pareto distribution with the shape parameter $\alpha > 0$ and the scale parameter $\gamma > 0$, that is, $P(\xi > x) = \frac{1}{(1+\gamma x)^\alpha},$ for $x \geq 0$. Then its pdf $g(x) = \frac{\alpha\gamma}{(1+\gamma x)^{\alpha+1}},\ x \geq 0$. Now it is easy to see that $\mathbb{E}[\xi] = [(\alpha-1)\gamma]^{-1}$. Furthermore, $\lambda_g = [(\alpha+1)\gamma]$ so that the maximum mean of the extracted exponential component equals $[(\alpha+1)\gamma]^{-1}$. In particular, the ratio of this extracted mean to the total mean equals $\frac{\alpha-1}{\alpha+1}$. It is independent of γ and increases from zero to 1 as α increases from 1 to ∞.

The example below illustrates the fact that there exist some distributions which may not have densities that belong to the family $\mathscr{H}$, but their components can be in $\mathscr{H}$ (with a scaling factor).

Example 4.4 (Weibull) Suppose ξ has a *Weibull*(α,γ) distribution with the shape parameter $\alpha < 1$ and scale parameter $\gamma > 0$. Its tail distribution $P(\xi > x) = \exp(-(\gamma x)^{\alpha})$, $x > 0$ and if g is pdf of ξ, then we have $\frac{-g'(x)}{g(x)} = \alpha\gamma^{\alpha}\frac{1}{x^{1-\alpha}} + \frac{(1-\alpha)}{x}$. Thus $\lambda_g = \sup_{y\in(0,\infty)}\left(-\frac{g'(y)}{g(y)}\right) = \infty$ and it is achieved at $y = 0$. Therefore, ξ can not be decomposed in the form (4.1). However, for any $a > 0$, $\sup_{y\in(a,\infty)}\left(-\frac{g'(y)}{g(y)}\right) < \infty$. Hence by fixing $a > 0$ and by letting $f(x) = cg(x)$ for $x \geq a$, and $f(x) = 0$ for $x < a$, we have $f \in \mathscr{H}$ and $g \geq f/c$, where $c = \frac{1}{P(\xi > a)}$.

The above example motivates a more general framework for extracting an exponential component. To see this, suppose that G is the distribution of a real valued random variable ξ such that for some $q \in (0,1]$ and $f \in \mathscr{H}$ and qf is a component of G, that is, $G(x_2) - G(x_1) \geq q\int_{x_1}^{x_2} f(y)dy$, $x_1 < x_2$. Let $\hat{\xi} \sim f$ and fix $\lambda \geq \lambda_f$, then from Theorem 4.1, $\hat{\xi} \stackrel{d}{=} E + Z$, where $E \sim Exp(\lambda)$, $Z \sim G_{\lambda}^{f}$ and they are independent. If $q = 1$ then f is the density of G, hence $\xi \stackrel{d}{=} \hat{\xi} \stackrel{d}{=} E + Z$. But if $q < 1$, then we can let $H(x) = \frac{G(x) - q\int_{-\infty}^{x} f(y)dy}{1-q}$, $x \in \mathbb{R}$. Clearly, H is a probability distribution function and $G(x) = (1-q)H(x) + q\int_{-\infty}^{x} f(y)dy$, $x \in \mathbb{R}$. In other words

$$\xi \stackrel{d}{=} (1-\beta)\tilde{\xi} + \beta(E+Z), \tag{4.2}$$

where β is a Bernoulli random variable with $P(\beta = 1) = q$, $\tilde{\xi} \sim H$ and $\tilde{\xi}$, E, β and Z are independent of each other.

5 MULTICLASS OPEN QUEUEING NETWORKS

A multiclass open queueing network can be characterized by d single server stations, K classes of customers, sequences of exogenous interarrival times $\{\xi_{k,n} : n \geq 1\}$ and service times $\{\eta_{k,n} : n \geq 1\}$ of each Class k and a probability matrix $P = (P_{kl})$ to specify the routing requirements of the customers, where P_{kl} is the probability of Class k customer becoming Class l customer upon completion of service at station $s(k)$ independently of all previous history; the customer exits the network with probability $1 - \sum_l P_{kl}$, where $s(k)$ is the station at which the Class k customers take service. If $\xi_{k,1} = \infty$, then we say that the external arrival process of Class k is *null*. Finally, L denotes the number of classes with *non-null* exogenous arrivals.

We assume that $(I - P')^{-1} = (I + P + P^2 + \dots)'$ exists, where P' is the transpose of the matrix P. Since (k, j)-element of $(I - P')^{-1}$ is the expected number of times a Class k customer visits Class l during its stay in the network, every customer who enter into the network will leave it eventually. Hence, the network described above is an open queueing network.

Like the $GI/G/1$ queue, any queueing network with a single non-null exogenous class regenerates at every instant when a customer arrival finds the system empty. Hence a regenerative structure trivially exists in these networks. So, without loss of generality we assume that $L \geq 2$ and first L classes are non-null exogenous.

Throughout the paper, we make the following assumptions on the network primitives.

(A1) $\xi_1, \xi_2, \dots, \xi_L, \eta_1, \eta_2, \dots, \eta_K$ are *i.i.d.* sequences and mutually independent, where $\xi_i = \{\xi_{i,n} : n \geq 1\}$ and $\eta_i = \{\eta_{i,n} : n \geq 1\}$.

(A2) There exists $p \geq 1$ such that $0 < \mathbb{E}[(\xi_{k,1})^p] < \infty$ for $k = 1, \dots, L$ and $0 < \mathbb{E}[(\eta_{k,1})^p] < \infty$ for $k = 1, \dots, K$.

(A3) For each $k = 2, \cdots, L$, there exists $f_k \in \mathscr{H}$ such that $P(\xi_{k,1} \in dx) \geq \bar{q}_k f_k(x)dx$ for some $\bar{q}_k > 0$, where $\mathscr{H}$ is defined in Section 4.

(A4) Distribution of $\xi_{1,1}$ is spreadout, that is, $P(\xi_{1,1} + \dots + \xi_{1,j} \in dx) \geq \bar{q}(x)dx$ for some non-negative function $\bar{q}(x)$ and integer j such that $\int_0^{\infty} \bar{q}(x)dx > 0$.

(A5) $P(\xi_{1,1} \geq x$ and $\sum_{l=1}^{K} \eta_{l,1} < x) > 0$ for some x.

Assumption (A1) is standard. It is clear from Section 4 that (A3) holds only when the interarrival times have either exponential or heavier tails. (A3) and (A4) are useful in establishing the ergodicity of the network as well as identifying regenerative structures in it, while (A2) and (A5) are useful in establishing finite moments of regeneration intervals as we see in the later sections. Notice that (A5) trivially holds when the interarrival times of Class 1 are unbounded or the service time distribution of every class has support in every neighborhood of 0.

We denote the common distribution of the interarrival times of Class k by F_k with the average arrival rate $\alpha_k := \frac{1}{\mathbb{E}[\xi_{k,1}]}$, where $k \in \{1,2,\dots L\}$. Similarly, the common distribution of the service times of Class k is denoted by H_k with the average service rate $\mu_k := \frac{1}{\mathbb{E}[\eta_{k,1}]}$, where $k \in \{1,2,\dots K\}$. Let $\mathscr{C}_i := \{k : s(k) = i\}$ be the *constituency* for station $i \in \{1,2,\cdots,d\}$. By letting $\sigma := (I-P')^{-1}\alpha$, one can interpret σ_k as the *effective* arrival rate to Class k. Then $\rho_i := \sum_{k\in\mathscr{C}_i} \frac{\sigma_k}{\mu_k}$ is the *nominal load* for server $i \in \{1,2,\dots,d\}$ per unit time. From (A3) and (4.2), without loss of generality, we can write

$$\xi_{k,n} = (1-\beta_{k,n})\tilde{\tilde{\xi}}_{k,n} + \beta_{k,n}(E_{k,n}+Z_{k,n}),\ k = 2,\dots,L,\ n \in \mathbb{Z}_+, \tag{5.1}$$

where $E_{k,n} \sim Exp(\lambda_k)$ for some $\lambda_k \geq \lambda_{f_k}$, $\beta_{k,n}$ is a Bernoulli random variable with $P(\beta_{k,n} = 1) = \bar{q}_k$, $Z_{k,n} \sim G^{f_k}_{\lambda_k}$ and $\tilde{\xi}_{k,n} \sim \dfrac{F_k(x) - \bar{q}_k \int_0^x f_k(y)dy}{1-\bar{q}_k}$, $x \geq 0$; and $\{\tilde{\xi}_{k,n} : n \geq\ 0\}$, $\{E_{k,n} : n \geq 0\}$, $\{\beta_{k,n} : n \geq 0\}$ and $\{Z_{k,n} : n \geq 0\}$ are *i.i.d.* sequences and independent of each other.

Now, to describe the network, we propose the following Markov process that splits the interarrival times into two components:

$$Y(t) \quad = \quad (Q(t),U(t),V(t)), \tag{5.2}$$

where $Q(t) = [Q_1(t),Q_2(t),\cdots,Q_K(t)]' \in \mathbb{Z}^K_+$. The process $Q_k(t)$ captures the number of Class k customers in the network at time t or, more generally, it can capture positions of every Class k customer present at station $s(k)$ (in the later case $Q_k(t)$ is an infinite dimensional vector). Hereafter, the notations $|Q_k(t)|$ and $|Q(t)|$ denotes, respectively, the number of Class k customers and the total number of customers present in the network at time t. The vector valued process $V(t) = [V_1(t),\dots,V_K(t)]' \in \mathbb{R}^K_+$ with $V_k(t)$ being the residual service time for Class k customer that is under service. We take $V_k(t) = 0$ whenever $|Q_k(t)| = 0$. The vector valued process $U(t) = [U_1(t),\ U_2^{(e)}(t),U_2^{(ne)}(t),\ \dots,U_L^{(e)}(t),U_L^{(ne)}(t)]' \in \mathbb{R}^{2L-1}_+$ such that $U_1(t)$ being the remaining time until the next Class 1 customer arrival, and at each instant of a Class $k \geq 2$ customer arrival, exponential and non-exponential components of the next interarrival time are generated independently and captured by $U_k^{(e)}$ and $U_k^{(ne)}$, respectively. Without loss of generality, we can assume that the non-exponential clock $U_k^{(ne)}$ decreases first linearly with rate 1 while exponential clock $U_k^{(e)}$ stays at the same value until $U_k^{(ne)}$ reaches zero. Thereafter, $U_k^{(e)}$ decreases linearly with rate 1 while other clock stays at zero. Next customer arrival happens when both the clocks are zero. Clearly, at any time t, $U_k^{(ne)}(t)+U_k^{(e)}(t)$ is the remaining time until the next Class k customer arrival. We say that Class k is in exponential phase at time t if $U_k^{(ne)}(t) = 0$. In Section 6.1, we see that this decomposition of the interarrival times play a crucial role in the construction of regenerations.

Let $\mathscr{Y}$ be the *state space* of the process Y and it is adapted to the filtration $\{\mathscr{F}_t : t \geq 0\}$ that is larger or equal to the natural filtration of the process Y. Hereafter, we assume that the state space $\mathscr{Y}$ is a complete and separable metric space with the *norm* defined by $|y| = |q| + |u| + |v|,\ y = (q,u,v) \in \mathscr{Y}$, where $|q|$ is the total queue length, $|u| = u_1 + \sum_{k=2}^L \left(u_k^{(e)} + u_k^{(ne)}\right)$ and $|v| = \sum_{k=1}^K v_k$.

Throughout the paper we assume that server at each station is busy whenever there is work to be done (work-conserving) and it stays idle whenever there is no work. Similar to Proposition 2.1 in Dai (1995), we can establish the strong Markov property of Y for a wide class of queueing disciplines, such as, FIFO (First-In-First-Out), LIFO (Last-In-First-Out), priority discipline, processor sharing, etc..

6 REGENERATIVE SIMULATION OF MULTICLASS NETWORKS

In this section, we identify a sequence of regeneration times and establish required finite moments on the associated regeneration intervals that satisfy the joint CLT.

6.1 Regenerations

For any set $B \in \mathscr{B}_{\mathscr{Y}}$, define the *first hitting time* $\tau_B := \inf\{t \geq 0 : Y(t) \in B\}$, the *first hitting time past* δ, $\tau_B(\delta) := \inf\{t > \delta : Y(t) \in B\}$ and the *first visiting time* $\Gamma_B := \inf\{t > \tau_{B^c} : Y(t) \in B\}$.

Let $D := \left\{(\vec{\mathbf{0}}, u, \vec{\mathbf{0}}) \in \mathscr{Y} : u_1 < \min_{k\geq 2}\left\{u_k^{(e)}\right\}, u_l^{(ne)} = 0, l \in \{2,\ldots,L\}\right\}$, where $\vec{\mathbf{0}}$ is all zeros vector of dimension K. The set D captures the states that Y can take when the network is empty and the next transition is triggered by a Class 1 customer arrival while all the other non-null exogenous classes are in exponential phase. Also let $S_{-1} = 0$ and the $(n+1)^{th}$ revisit instant on set D, $S_n = \theta_{S_{n-1}} \circ \Gamma_D$, where θ is the shift operator on the sample paths of the process Y.

Let $\tilde{D} := \left\{(q,u,v) \in \mathscr{Y} : |q_1| = 1, |q_k| = 0, \forall\, k \in \{2,\ldots,K\}, u_l^{(ne)} = 0, \forall\, l \in \{2,\ldots,L\}\right\}$ and $T_n = S_n + U_1(S_n)$ be the first instant when an arrival of Class 1 finds the system empty past S_n. Then for each $n \geq 0$, $Y(T_n-) \in D$ and $Y(T_n) \in \tilde{D}$. Consider $A(r,s) = \left\{y \in \tilde{D} : u_1 \leq r_1,\ v_1 \leq s, u_k^{(e)} \leq r_k, \forall\, k \in \{2,\ldots,L\}\right\}$, for $r = (r_1,\ldots,r_L) \in \mathbb{R}_+^L, s \in \mathbb{R}_+$, and let φ be the probability measure on $\mathscr{B}_{\mathscr{Y}}$ such that $\varphi(A(r,s)) = F_1(r_1)H_1(s)\prod_{k=2}^{L}\left(1 - e^{-\lambda_k r_k}\right)$, where F_k, H_k are the distributions of Class k interarrival and service times (refer to Section 5). Then it is clear that $\varphi\left(\tilde{D}\right) = 1$.

One can easily check that when $Y(0) \sim \varphi$, the process Y is a non-delayed regenerative process with regeneration times $\{T_n : n \geq 0\}$ and $Y(T_n) \sim \varphi$ for every $n \geq 0$.

6.2 Moments of Regeneration Intervals

In this section, we establish the finite p^{th} moments of the regeneration intervals, where p is the parameter used in (A2) that guarantees finite p^{th} moments of the interarrival and service times. As mentioned earlier, proofs are provided in Moka and Juneja (2013). Hereafter, we make the following assumption on the Markov process.

(A6) There exists $t_0 > 1$ such that $\lim_{|y|\to\infty} \dfrac{1}{|y|^p}\mathbb{E}_y\left[|Y(t|y|)|^p\right] = 0, \ \ \forall\, t \geq t_0.$

Remark 6.1 Dai (1995) shows that, under mild conditions, (A6) holds when Assumptions (A1) and (A2) hold and the fluid model of the network is stable. In particular, Dai (1995) considers some important networks like re-entrant lines and generalized Jackson networks, and a wide variety of queueing disciplines; and shows that the fluid model is stable if the nominal traffic condition (that is, $\rho_i < 1$, for each $i = 1,\ldots,d$) holds. Recent work of Schönlein and Wirth (2012) shows that under certain conditions, fluid model is stable if and only if there exists a Lyapunov function (also refer to Schönlein (2012)).

Now we show that the p^{th} moments of the regeneration intervals are finite. Let $C_s := \{y \in \mathscr{Y} : |y| \leq s\}$, for $s \geq 0$. Lemma 6.1 below establishes that for any initial state, when the p^{th} moments of the interarrival and service times are finite, the time required for the queue length to become smaller than a certain level has finite p^{th} moment. Furthermore, it says that there exists a bounded set such that the process visits the set infinitely often and the associated intervals have finite p^{th} moments. The proof of Lemma 6.1 mainly depends on (A6).

Lemma 6.1 *There exist constants* $c_1,\ c_2,\ s_0 > 0$ *such that* $\mathbb{E}_y\left[(\tau_{C_s}(\delta))^p\right] \leq c_1 + c_2|y|^p,\ \ y \in \mathscr{Y},\ s \geq s_0,\ \delta > 0$*, Furthermore, for any bounded set A,* $\sup_{y\in A}\mathbb{E}_y\left[(\tau_{C_s}(\delta))^p\right] < \infty,\ \ s \geq s_0$.

Since the networks are open, there is a positive probability of any customer leaving the network by making at most K distinct transitions, where the transition refers to the customer's class change upon completion of it's service. From (A5), it is clear that with a positive probability any Class 1 customer who arrived into empty system can leave the network before the next arrival of the same class when there are no arrivals from other classes. Using this fact, we can show that there exists $\delta > 0$ such that $\inf_{y \in C_s} P_y(\Gamma_D + U_1(\Gamma_D) \leq \delta) > 0$. Then using Lemma 6.1 and the standard geometric argument, we can establish the following result.

Lemma 6.2 *There exist constants $c_3,\ c_4 > 0$ such that $\mathbb{E}_y\left[(\Gamma_D + U_1(\Gamma_D))^p\right] < c_3 + c_4|y|^p,\ y \in \mathscr{Y}$.*

From Lemma 6.1 and Lemma 6.2, we can argue that the first cycle length T_0 is a proper random variable for any initial state $y \in \mathscr{Y}$. Hence, hereafter, we assume that the regenerative process Y is non-delayed (that is, $T_0 = 0$) and hence the initial state $Y(0) \sim \varphi$. The proposition below establishes finite moments of the regeneration intervals, $\tau_n = T_n - T_{n-1},\ n \geq 1$, using the fact that $\mathbb{E}\left[Y(0)^p\right] < \infty$. The non-lattice property of the distribution of τ_1 simply follows from the spreadout Assumption (A4) on interarrival times of Class 1.

Proposition 6.1 *The distribution of τ_1 is non-lattice and $\mathbb{E}_\varphi[\tau_1^p] < \infty$.*

6.3 Moments of $\mathbf{R_1}$

Let h be a non-negative real valued function defined on $\mathscr{Y}$ and $R_i = \int_{T_{i-1}}^{T_i} h(Y(s))\,ds$, for $i \geq 1$. When h is bounded (e.g., $h(x) = I(x > 10)$), there exists a constant c such that $h < c$ and $\mathbb{E}_\varphi\left[R_1^p\right] = \mathbb{E}_\varphi\left[\left(\int_0^{\tau_1} h(Y(t))dt\right)^p\right] \leq c^p \mathbb{E}_\varphi\left[\tau_1^p\right] < \infty$. But, when h is unbounded, $\mathbb{E}_\varphi\left[R_1^p\right]$ can be infinite even though $\mathbb{E}_\varphi[\tau_1^p] < \infty$. However, we can guarantee these moments even when h is unbounded under some additional conditions as shown in the following proposition.

For each $r > 0$ and $s \geq s_0$, define $J_{r,s}(y) := \mathbb{E}_y\left[\left(\int_0^{\tau_{C_s}(\delta)} h(Y(t))dt\right)^r\right],\ y \in \mathscr{Y}$, where δ is given by Lemma 6.2 and s_0 is given by Lemma 6.1.

Proposition 6.2 *Suppose that for a given $r > 0$, there exists $s \geq s_0$ such that $J_{r,s}(\cdot)$ is uniformly bounded on C_s and $\mathbb{E}\left[J_{r,s}(Y(0))\right] < \infty$. Then $\mathbb{E}_\varphi\left[R_1^r\right] < \infty$.*

The following example illustrates one possible application of Proposition 6.2.

Example 6.1 Suppose that $h(Y(t)) \leq |Y(t)|$, where $|\cdot|$ denotes the norm on the state space $\mathscr{Y}$ (refer to Section 5). For example, if goal is to estimate the steady-state expected number of customers in the network then $h(Y(t)) = |Q(t)| \leq |Y(t)|$.

Now we show that if $p \geq 9$ in (A2) then $\mathbb{E}_\varphi\left[R_1^4\right] < \infty$ (this is needed for Theorem 3.2 (iii) to hold). To see this, observe from Lemma 6.1 that $\mathbb{E}_\varphi\left[\tau_{C_s}(\delta)^8\right] \leq c_1 + c_2\mathbb{E}\left[|Y(0)|^8\right] < \infty$ for any $s > s_0$. Similar to Proposition 5.3 of Dai and Meyn (1995), it can be shown that there exist $s \geq s_0$ and $c < \infty$ such that $\mathbb{E}_y\left[\int_0^{\tau_{C_s}(\delta)} |Y(t)|^p dt\right] \leq c(|y|^{p+1} + 1)$. Since $p \geq 9$, we have that $\mathbb{E}\left[|Y(0)|^9\right] < \infty$, and hence $\mathbb{E}_\varphi\left[\int_0^{\tau_{C_s}(\delta)} |Y(t)|^8 dt\right] \leq c\left(\mathbb{E}\left[|Y(0)|^9\right] + 1\right) < \infty$. Using Cauchy-Schwarz inequality, Jensen's inequality and the fact that $\tau_{C_s}(\delta) \geq 1$, we can show that $\mathbb{E}\left[J_{4,s}(Y(0))\right] \leq \left(\mathbb{E}_\varphi\left[\tau_{C_s}(\delta)^8\right]\mathbb{E}_\varphi\left[\int_0^{\tau_{C_s}(\delta)} |Y(t)|^8 dt\right]\right)^{\frac{1}{2}} < \infty$. Similarly, uniform boundedness of $J_{4,s}(y)$ on C_s can be established and from Proposition 6.2, it follows that $\mathbb{E}_\varphi\left[R_1^q\right] < \infty$.

7 CHOICE OF OPTIMAL REGENERATIVE STRUCTURE

In this section, we show that under certain assumptions, if one regenerative structure is a subsequence of another then it is optimal to choose the original sequence over the subsequence (optimal in the sense that the AVSDE associated with the original sequence is smaller than that associated with the subsequence). Using this result, we show that the selection of the interarrival time decomposition with largest mean exponential component minimizes the AVSDE.

Suppose that X is a non-delayed regenerative process with the regeneration times $\{T_n : n \geq 0\}$ and the initial state distribution φ. Let, for each $i \geq 1$, $\tau_i = T_i - T_{i-1}$, $R_i = \int_{T_{i-1}}^{T_i} h(X(s))\,ds$, where h is a non-negative, real valued function. Assume that there exists a filtration $\mathscr{G} := \{\mathscr{G}_n, n \geq 0\}$ and a strictly increasing sequence of integer valued stopping times $0 = \nu_0 < \nu_1 < \cdots$ adapted to $\mathscr{G}$ with $\nu_n - \nu_{n-1} \stackrel{d}{=} \nu_1,\ n \geq 1$ such that **(i)** $\{(R_i, \tau_i) : i \geq 1\}$ is adapted to $\mathscr{G}$, **(ii)** $\{(R_i, \tau_i) : i \geq n\}$ is independent of $\mathscr{G}_{n-1}$ for all $n \geq 1$ and **(iii)** the sequence $\{S_n = T_{\nu_n} : n \geq 0\}$ is another regenerative structure of X.

Proposition 7.1 *Under the above setup, if* $\mathbb{E}_\varphi\left[\left(\int_0^{S_1}[h(X(s))+1]\,ds\right)^4\right] < \infty$*, then the AVSDE associated with* $\{S_n : n \geq 0\}$ *is at least as large as that associated with* $\{T_n : n \geq 0\}$.

Andradóttir, Calvin, and Glynn (1995) establish a similar result on classical regenerative processes and consider stopping times that depend only on history of the process, and hence these stopping times turn out to be geometrically distributed due to the $i.i.d.$ nature of classical regenerations. Our result is a mild extension of Theorem 1 in Andradóttir, Calvin, and Glynn (1995) in the sense that the stopping time can be more general (hence, its distribution may not be geometric).

Now using Proposition 7.1 and path-wise construction, we show that the selection of the interarrival time decomposition with largest mean exponential component (that is, $\lambda_k = \lambda_{f_k}$) minimizes the AVSDE. Assume that the function h satisfies $h(Y(t)) = \hat{h}(Q(t), V(t))$ (a well known example is when $h(Y(t)) = |Q(t)|$ to estimate the steady-state expected customers in the network). Recall that the arrival clock of each Class $k \in \{2, \ldots, L\}$ is updated based on the decomposition (5.1), that is,

$$\xi_{k,n} = (1-\beta_{k,n})\tilde{\xi}_{k,n} + \beta_{k,n}(E_{k,n} + Z_{k,n}),\ n \in \mathbb{Z}_+,$$

where $\xi_{k,n}$ is the n^{th} interarrival time of Class k, $E_{k,n} \sim Exp(\lambda_k)$ for some $\lambda_k \geq \lambda_{f_k}$, $\beta_{k,n}$ is a Bernoulli random variable with $P(\beta_{k,n} = 1) = \bar{q}_k$, $Z_{k,n} \sim G_{\lambda_k}^{f_k}$ and $\tilde{\xi}_{k,n} \sim \dfrac{F_k(x) - \bar{q}_k \int_0^x f_k(y)dy}{1-\bar{q}_k}$, $x \geq 0$. Note that $\{\tilde{\xi}_{k,n} : n \geq 0\}$, $\{E_{k,n} : n \geq 0\}$, $\{\beta_{k,n} : n \geq 0\}$ and $\{Z_{k,n} : n \geq 0\}$ are $i.i.d.$ sequences and independent of each other.

Suppose that $\lambda_k > \lambda_{f_k}$. It is well known that when $E_{k,n} \sim Exp(\lambda_{f_k})$ and $E'_{k,n} \sim Exp(\lambda_k)$, with out loss of generality, one can write $E_{k,n} = E'_{k,n} + Z'_{k,n},\ n \geq 1$ for an $i.i.d.$ sequence of almost surely positive random variables $Z'_{k,n},\ n \geq 1$ which are independent of $E'_{k,n},\ n \geq 1$. Since $E_{k,n} > E'_{k,n}$ for all $n \geq 0$, the regenerative structure associated with exponential clock rate λ_k (denote it by $\{S_n : n \geq 0\}$) is a subsequence of that associated with exponential clock rate λ_{f_k} (denote it by $\{T_n : n \geq 0\}$). By letting $\tau_i = T_i - T_{i-1}$ and $R_i = \int_{T_{i-1}}^{T_i} h(Y(s))\,ds$ for $i \geq 1$, we can easily see that $\{(R_i, \tau_i) : i \geq 1\}$ is adapted to $\mathscr{G} = \{\mathscr{G}_n = \sigma((R_i, \tau_i), \mathbf{1}_i)\}$ and $\{(R_i, \tau_i) : i \geq n\}$ is independent of $\mathscr{G}_{n-1}$ for all $n \geq 1$, where $\mathbf{1}_i = \mathbb{I}(T_i = S_j$ for some $j)$. If we assume that $\mathbb{E}_\varphi\left[\left(\int_0^{S_1}[h(X(s))+1]\right)^4\right] < \infty$ (for example, if $h(Y(t)) \leq |Y(t)|$ then, from Example 6.1, this is guaranteed when the 9^{th} moments of the interarrival and service times are finite), then from Proposition 7.1, it is optimal to choose $\lambda_k = \lambda_{f_k}$ for each $k = 2, 3, \ldots, L$.

8 SIMULATION

In this section, we present one preliminary numerical example to illustrate the regenerative simulation method proposed in the previous sections. Consider a multiclass system with four stations and populated with four classes of customers. There are three non-null exogenous classes, namely Class 1 to Class 3, which get service at three different stations, namely station 1 to station 3, respectively, and after the service completion they will become Class 4 customers to get service at station 4. Every customer who get service at station 4 will leave the system. Every station follows work-conserving FCFS service discipline. The network primitives are specified as follows: The interarrival time distributions of the non-null exogenous classes 1 to 3 are $Uniform(0,40)$, $Pareto(10, 1/18)$ and $Pareto(10, 1/9)$, respectively. The service time distributions of the classes 1 to 4, are $\frac{3}{4}Exp(3/20) + \frac{1}{4}Exp(1/20)$, $\frac{1}{2}Exp(2/3) + \frac{1}{2}Exp(2)$, $Exp(4/3)$ and $Exp(5)$, respectively.

Here, our goal is to estimate the steady state expected number of customers in the system (that is, $h(Y(t)) = |Q(t)|$). Assumptions (A1) to (A5) are trivially satisfied. This network is a special case of generalized Jackson networks and hence the fluid model is stable as the nominal load $\rho_i < 1,\ i = 1,\ldots,4$ (see Remark 6.1) and thus Assumption (A6) holds. Notice that the 9^{th} moments of the interarrival and service times are finite. Hence, from Example 6.1, $\mathbb{E}_\varphi\left[R_1^4 + \tau_1^4\right] < \infty$. This guarantees asymptotically valid confidence intervals and also the finiteness of the AVSDE.

We generate the underlying Markov process using the above specifications, identify the regenerations associated with all the three scenarios along the same process and estimate the required parameters corresponding to each scenario. Table 1 displays the simulation results. It can be observed that the estimated AVSDE is increasing as the rates $\lambda_k, k = 2,3$, of exponential components are increasing and is small when $\lambda_k = \lambda_{f_k},\ k = 2,3$. As expected, the estimated TAVC is not changing with λ_k.

Table 1: Estimation of steady state expected number of customers in a system with four stations and four classes. The total duration of the simulation is 10^7 units of time.

Estimate	$\lambda_k = \lambda_{f_k}$ for $k = 2,3$	$\lambda_k = 1.5\lambda_{f_k}$ for $k = 2,3$	$\lambda_k = 2\lambda_{f_k}$ for $k = 2,3$
No. of cycles generated	13144	3892	1663
Estimated steady state mean	5.86	5.86	5.86
95% confidence interval	$5.86 \pm 6.2 \times 10^{-6}$	$5.86 \pm 6.2 \times 10^{-6}$	$5.86 \pm 6.3 \times 10^{-6}$
Estimated TAVC	1.0×10^3	1.0×10^3	1.0×10^3
Estimated AVSDE	1.8×10^6	4.6×10^6	6×10^6

9 SUPEREXPONENTIAL CASE

(A3) restricts the above proposed regenerative simulation to multiclass networks where the interarrival times have either exponential or heavier tail distributions. In this section, we propose a regenerative simulation method which can be implemented even when the interarrival times of some (or all) classes have superexponential continuous distributions. Without loss of generality, one can assume that the network of interest is not satisfying both (A3) and (A5) together. Then we replace these assumptions, respectively, with the following assumptions:

(A3′) For each Class $k = 1,\ldots,L$, there exist a density g_k and a constant $\bar{q}_k > 0$ such that either
(i) $g_k \in \mathscr{H}$ and $P(\xi_{k,1} \in dx) \geq \bar{q}_k g_k(x)dx$, where $\mathscr{H}$ is defined in Section 4, or
(ii) $\xi_{k,1} \sim f_k$ and there exist a_k, b_k $(a_k < b_k)$ such that $\frac{f_k(y+x)}{1-F_k(y)} \ \geq \ \bar{q}_k \ g_k(x),\ x \geq 0,\ \ y \in [a_k, b_k]$.

(A5′) Either all the non-null exogenous classes have unbounded interarrival times or $P(\sum_{l=1}^K \eta_{l,1} \leq x) > 0$ for every $x > 0$.

Assumption (A3′) is important for identifying regenerations in the network while (A5′) is crucial for establishing the moments of the regeneration intervals. Let $\mathscr{N}$ (respectively, $\mathscr{M}$) be the set of classes that satisfy condition (i) (respectively, condition (ii)) in (A3′). From previous sections, it is clear that the interarrival time distribution of each class in $\mathscr{N}$ has an exponential or heavier tail while the interarrival time distributions of classes in $\mathscr{M}$ can even have superexponential distributions. Notice that in addition to the above two assumptions, we still assume (A1), (A2), (A4) and (A6). Now we see some examples of superexponential distributions that satisfy (A3′).

Example 9.1 (**Uniform**) Suppose that f and g are densities of uniform distributions with support $[c,d]$ and $\left[0, \frac{d-c}{2}\right]$, respectively, for some $0 < c < d$. It is easy to see that, for each $y \in [c,d]$, $\frac{f(y+x)}{1-F(y)} = \frac{1}{d-y}$ for $x \in [y,d]$, otherwise, zero, where F is the distribution associated with the density f. Hence, we note that $\frac{f(y+x)}{1-F(y)} \geq \frac{1}{2}g(x)$ for $y \in [c, \frac{d+c}{2}]$ and $x \geq 0$; so (A3′) holds.

Similarly, one can show that folded normal distribution, and $Weibull(\alpha,\gamma)$ with shape parameter $\alpha > 1$ and scale parameter $\gamma > 0$ are examples of superexponential distributions that satisfy (A3′) with unbounded support.

To identify the regenerations, we describe the network with the strong Markov process $Y(t) = \big(Q(t), \tilde{U}(t), V(t)\big)$, where, as usual, Q and V are the queue length and remaining service time processes, respectively, and $\tilde{U}(t) = [\tilde{U}_1(t), \ldots, \tilde{U}_L(t)]'$. If $k \in \mathscr{M}$ then $\tilde{U}_k(t)$ is the remaining time until the next Class k customer arrival, else if $k \in \mathscr{N}$, $\tilde{U}_k(t) = \left[U_k^{(e)}(t), U_k^{(ne)}(t)\right]'$, where the exponential component $U_k^{(e)}$ and the non-exponential component $U_k^{(ne)}$ are defined in Section 5. During the simulation it is possible to keep track of age of the last arrival of each Class $k \in \mathscr{M}$, and hence for each $k \in \mathscr{M}$, let $A_k(t)$ denote the time passed since the last arrival of Class k.

Let S_n be the n^{th} instant when a Class 1 arrival finds the system empty, the age $A_k(t) \in [a_k, b_k]$ for each $k \in \mathscr{M}$ and all the classes in $\mathscr{N}$ are in exponential phase. Now we argue that S_n can be a regeneration instant with a positive probability. To see this, observe that for any Class $k \in \mathscr{M}$ if the age $A_k(t) \in [a_k, b_k]$ then the remaining time $R_k(t)$ can have density g_k (which is independent of the age $A_k(t)$) with probability q_k. Let β_k be a Bernoulli random variable with $P(\beta_k = 1) = q_k$ such that when $\beta_k = 1$, remaining time until the next arrival has density g_k. Then at every instant S_n and for each Class $k \in \mathscr{M}$, one can compute the posterior probabilities,

$$P(\beta_k = 1 \,|A_k(S_n) = y, R_k(S_n) = x) = \frac{q_k g_k(x)\,(1 - F_k(y))}{f_k(y+x)}. \tag{9.1}$$

The above equation denotes the probability that the remaining time sample $R_k(S_n)$ is generated from g_k given that the age $A_k(S_n) = y$ and the remaining time until the next arrival $R_k(S_n) = x$. These posterior probabilities can be obtained easily using Bayes' theorem. At every S_n and for every Class $k \in \mathscr{M}$, we can toss a coin with probability of head equal to (9.1). If all coins show up heads, then we declare a regeneration at S_n. Hence, we have identified an implementable regenerative structure for the process Y, denote it as $\{T_n : n \geq 0\}$. It is clear that the distribution φ of the process Y at regeneration instants is just the product measure generated by F_1, H_1, $Exp(\lambda_k)$, $k \in \mathscr{N}$ and g_l, $l \in \mathscr{M}$.

It is important to note that all the results in Sections 6.2 and 6.3 can be extended to present setup. In particular, we can establish the finite moments of regeneration times and show that Proposition 6.2 holds. Also one can prove that the selection of the interarrival time decomposition with largest mean exponential component, for every class in $\mathscr{N}$, results in the minimum AVSDE.

10 SUMMARY

In this paper, first we showed that, under mild conditions, a random variable with an exponential or heavier tail can be re-expressed as a mixture of sums of independent random variables where one of the constituents is exponential. Using this result, we developed an implementable regenerative simulation technique for multiclass open queueing networks where the interarrival times have exponential or heavier tails. Under mild stability conditions, we established that finite moments of the interarrival and service times are sufficient to guarantee asymptotically valid confidence intervals. We also showed that the selection of the interarrival time decomposition with largest mean exponential component minimizes the AVSDE. Finally, we discussed a different regenerative simulation technique that can be applicable even when the interarrival times of some (or all) classes have superexponential tails.

REFERENCES

Andradóttir, S., J. M. Calvin, and P. W. Glynn. 1995. "Accelerated Regeneration for Markov Chain Simulations". *Probability in the Engineering and Informational Sciences* 9:497–523.

Asmussen, S. 2003. *Applied probability and queues*. Second ed, Volume 51 of *Applications of Mathematics (New York)*. New York: Springer-Verlag. Stochastic Modelling and Applied Probability.

Athreya, K. B., and P. Ney. 1978. "A New Approach to the Limit Theory of Recurrent Markov Chains". *Transactions of the American Mathematical Society* 245:pp. 493–501.

Dai, J. G. 1995. "On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Limit Models". *Annals of Applied Probability* 5 (1): 49–77.

Dai, J. G., and S. Meyn. 1995, nov. "Stability and convergence of moments for multiclass queueing networks via fluid limit models". *Automatic Control, IEEE Transactions on* 40 (11): 1889 –1904.

Glynn, P. W. 2006. "Chapter 16 Simulation Algorithms for Regenerative Processes". In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Volume 13 of *Handbooks in Operations Research and Management Science*, 477–500. Elsevier.

Glynn, P. W., and D. L. Iglehart. 1987. "A Joint Central Limit Theorem for the Sample Mean and Regenerative Variance Estimator". *Annals of Operations Research* 8:41 – 55.

Henderson, S., and P. W. Glynn. 1999. "Can the regenerative method be applied to discrete-event simulation?". In *Simulation Conference Proceedings, 1999 Winter*, Volume 1, 367 –373 vol.1.

Henderson, S. G., and P. W. Glynn. 2001, October. "Regenerative steady-state simulation of discrete-event systems". *ACM Trans. Model. Comput. Simul.* 11 (4): 313–345.

Iglehart, D. L., and G. S. Shedler. 1979. "Regenerative simulation of response times in networks of queues with multiple job types". *Acta Informatica* 12:159–175.

Iglehart, D. L., and G. S. Shedler. 1980. *Regenerative simulation of response times in networks of queues*, Volume 26 of *Lecture Notes in Control and Information Sciences*. Berlin: Springer-Verlag.

Iglehart, D. L., and G. S. Shedler. 1983. "Statistical efficiency of regenerative simulation methods for networks of queues". *Adv. in Appl. Probab.* 15 (1): 183–197.

Li, J. 1997. "An Approximation Method for the Analysis of GI/G/1 Queues". *Operations Research* 45 (1): pp. 140–144.

Miller, G. K., and U. N. Bhat. 1997. "Estimation for renewal processes with unobservable gamma or Erlang interarrival times". *Journal of Statistical Planning and Inference* 61 (2): 355 – 372.

Moka, S. B. and Juneja, S. 2013. "Regenerative Simulation for Multiclass Open Queueing Networks". http://www.tcs.tifr.res.in/~sarathmoka/preprint.pdf.

Nummelin, E. 1978. "A splitting technique for Harris recurrent Markov chains". *Z. Wahrsch. Verw. Gebiete* 43 (4): 309–318.

Schönlein, M. 2012, September. "A Lyapunov view on positive Harris recurrence of multiclass queueing networks". *ArXiv e-prints*.

Schönlein, M., and F. Wirth. 2012. "On converse Lyapunov theorems for fluid network models". *Queueing Systems* 70:339–367.

Smith, W. L. 1955. "Regenerative stochastic processes". *Proc. Roy. Soc. London. Ser. A.* 232:6–31.

Thorisson, H. 1983. "The coupling of regenerative processes". *Adv. in Appl. Probab.* 15 (3): 531–561.

Thorisson, H. 2000. *Coupling, Stationarity, and Regeneration*. New York: Springer-Verlag.

AUTHOR BIOGRAPHIES

SARAT BABU MOKA is a PhD student in the School of Technology and Computer Science at Tata Institute of Fundamental Research, Mumbai, India. He received a Bachelor of Engineering from Andhra University and Master of Engineering from Indian Institute of Science, Bangalore. His research interests lie in applications of probability and game theory with a particular focus on queueing systems. His email address is sarathmoka@tcs.tifr.res.in.

SANDEEP JUNEJA is a Professor in the School of Technology and Computer Science at Tata Institute of Fundamental Research. His Ph.D. was from Department of Operations Research at Stanford University (1993). His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial board on Mathematics of Operations Research. His email address is juneja@tifr.res.in and his web page is http://www.tcs.tifr.res.in/~sandeepj.