Perfect Sampling for Loss Networks
Jing Dong

Following the definition in Loss Networks by F. Kelly (Annals of Applied Probability, (1):319?378, 1991), we consider a generalized loss network with $J$ stations, labeled $1, 2, \cdots, J$ and suppose that station $j$ comprises $C_j$ servers. We have $L$ possible routes, labeled $1, 2, \ldots, L$ and for each route $l$, a $J$ dimensional routing vector $P_l$. $P_l$ is consist of 1’s and 0’s, where $P_l(j) = 1$ means route $l$ requires a server at station $j$. A routing request $l$ is blocked and thus lost if any station $j$ with $P_l(j) = 1$ is full at the arrival time of the request. Customers requesting route $l$ form a renewal process with i.i.d. interarrival times $\{X_n^{(l)} : n \geq 1\}$. The CDF of $X_n^{(l)}$ is $G_l$. Independent of the arrival process, the service times $\{V_n^{(l)} : n \geq 1\}$ are also i.i.d. with CDF $F_l$. We assume that $G_l$’s and $F_l$’s satisfy that they all have finite first moments.

Following the same strategy as in the many-server loss queue case, we first couple the loss network with a network of infinite-server stations. Notice that no customer is blocked or lost in the infinite server system, thus it imposes an upper bound on the number of jobs in the loss system. Let $Q_j(t, y)$ denote the number of jobs in the $j$-th station with remaining service time strictly greater than $y$ at time $t$. Note that a class $l$ job with remaining service time greater than $y$ in the system will be counted in all $Q_j(t, y)$’s with $P_l(j) = 1$. Let $R_j(t)$ denote the longest remaining service time among all customers in station $j$ at time $t$. Let $R(t) = \max_{1 \leq j \leq J} \{R_j(t)\}$. Then similar to the many server loss queue, we define a random time $\tau'$ satisfying the following conditions:

1) $R(\tau') \leq |\tau'|,$

2) $\inf_{\tau' \leq t' \leq R(\tau')} \inf_{1 \leq j \leq J} \{C_j - Q_j(t, 0)\} \geq 0,$

i.e. all links are operating below capacity on the interval $[\tau', \tau' + R(t')]$.

At time $\tau' + R(\tau')$, everyone in the network of infinite-server stations will be in the loss network as well. Thus from then on (forwards in time), we can update the loss system using the inputs of the infinite-server system.

In order to simulate the network of infinite-server stations with $L$ types of routing requests, we simulate $L$ independent networks of infinite-server stations; each dealing with a single type of routing request. Then we do a superposition of them. The simulation of each independent network of infinite-server stations are exactly the same as what we have described in Section 2.3, as a type $l$ routing request occupies a server from each station $j$ with $P_l(j) = 1$ simultaneously and for the same amount of time. For the $l$-th system, let $Z_n^{(l)} = (A_n^{(l)}, V_n^{(l)})$ represent the arrival time and service time of the $n$-th
routing request counting backwards in time and $\kappa(l)$ be a random time satisfying that $V_n(l) \leq |A_n(l)|$ for all $n \geq \kappa(l)$. Then following the procedure described in Section 2.3, we will be able to simulate $\kappa(l)$ as the maximum of two random times associated the arrival process and service time process respectively.

We now consider a sequence of systems indexed by $s \in \mathbb{N}^+$. We speed up the the arrival rate of the $s$-th system by $s$, i.e. $X_{n(s)}(l) = X_n(l)/s$, and keep the service rate fixed.

**Theorem 1.** Assume $EX_n^{(l)} < \infty$ (C).

1. if $E[(V_n(l))^q] < \infty$ for some $q > 2$, then
   $$E_s^{(l)} \kappa(l) = O(s^{q/(q-1)});$$

2. if we further assume $E[\exp(\theta V_n(l))] < \infty$ for some $\theta > 0$, then
   $$E_s^{(l)} \kappa(l) = O(s \log s)$$

for $l = 1, 2, \cdots, L$.

In what follows we shall hold the number of routing request types, $L$, fixed. We run $L$ independent networks of infinite-server stations. Network $l$ serves routing request of type $l$ only, for $l = 1, 2, \cdots, L$. Let $Q(l)(t, 0)$ denote the number of jobs in network $l$ at time $t$ and $R(l)(t)$ denote the maximum remaining service time among all jobs in the network at time $t$. Then we have $R(t) = \max\{R(l)(t) : 1 \leq l \leq L\}$.

We consider two asymptotic regimes. One is the QD regime where for the base system we have

$$\sum_{l=1}^{L} \frac{EV_n^{(l)}}{EX_n^{(l)}} P_l(l) < C_j. \quad (1)$$

For the $s$-th system, the number of servers in the $j$-th station is $C_j(s) = sC_j$ for $j = 1, 2, \cdots, J$.

Assign a fixed number $H_l$ to each route $l$. $H_l$ is well chosen such that $E[V_n(l)]/EX_n^{(l)} < H_l$ and $\sum_{l=1}^{L} H_l P_l(j) \leq C_j$. This is doable because of (1). Let $H_l^s = sH_l$. Define a random time $\bar{\tau}'$ satisfying the following two conditions:

1. $R(l)(\bar{\tau}') \leq |\bar{\tau}'|$ for $l = 1, 2, \cdots, L$,
2. $\inf_{\bar{\tau}' \leq t \leq \bar{\tau}' + R(l)} \{H_l - Q_l(t, 0)\} \geq 0$ for $l = 1, 2, \cdots, L$.

Notice that $\bar{\tau}'$ is an upper bound on $\tau'$.

**Theorem 2.** Assume $EX_n^{(l)} < \infty$ and $X_n^{(l)}$’s are non-lattice and strictly positive. We also assume $E[(V_n(l))^q] < \infty$ for any $q > 0$ and $F_l$ is continuous. Then

$$E_s^{(l)} \bar{\tau}' = o(s^\delta)$$

for any $\delta > 0$.  

2
The other asymptotic regime is the QED regime where for the base system we have

$$\sum_{i=1}^{L} \frac{EV_n^{(l)}}{EX_n^{(l)}} P_j(l) = C_j$$

and the number of servers in the $j$-th station of the $s$-th system is $C_s^j = sC_j + \beta_j \sqrt{s}$ for $j = 1, 2, \cdots, J$

We then let $I_l = E[V_n^{(l)}]/E[X_n^{(l)}]$ and $I_l^s = sI_l + a_l \sqrt{s}$ where $a_l$'s are well chosen such that $\sum_{l=1}^{L} a_l P_j(l) \leq \beta_j$.

We define a random time $\tilde{\tau}'$ that satisfies the following two conditions:

1) $R^{(l)}(\tilde{\tau}') \leq |\tilde{\tau}'|$ for $l = 1, 2, \cdots, L$,

2) $\inf_{t \leq \tilde{\tau}' + R(\tilde{\tau}')} \{ I_l - Q^{(l)}(t, 0) \} \geq 0$ for $l = 1, 2, \cdots, L$.

As before, $\tilde{\tau}'$ is an upper bound on $\tau'$.

**Theorem 3.** Assume $E[(X_n^{(l)})^2] < \infty$. We also assume $E[(V_n^{(l)})^q] < \infty$ for any $q > 0$. Then for $b_j$'s large enough, we have

$$\log E_n^s \tau' = o(s^\delta)$$

for any $\delta > 0$. 