Busy Period Length and Higher level First passage probabilities

Chaitanya Garikiparthi, Appie van de Liefvoort, Kenneth Mitchell

Abstract

I. Introduction

II. Conditional density for the $min(A, S)$ process

Consider two contesting processes, A and S (representing Arrivals and Service completions), both represented by the corrosponding matrix exponential notations $< \boldsymbol{p}_a, \boldsymbol{B}_a, \boldsymbol{e}_a >$ and $< \boldsymbol{p}_s, \boldsymbol{B}_s, \boldsymbol{e}_s >$. Then the conditional density for the min process given that the arrival process occurs before the service process is,

$$
\begin{aligned}
\text{Pr[} min(A, S) = t \mid A < S \text{]} &= \frac{\text{Pr[} min(A, S) = t \text{ and } A < S \text{]}}{\text{Pr[} A < S \text{]}} \\
&= \frac{\boldsymbol{p}_a \exp(-\boldsymbol{B}_a t)\boldsymbol{L}_a \boldsymbol{e}_a \boldsymbol{p}_s \exp(-\boldsymbol{B}_s t)\boldsymbol{e}_s}{\boldsymbol{p}_a \widehat{\boldsymbol{p}_s}(\widehat{\boldsymbol{B}_a} + \widehat{\boldsymbol{B}_s})^{-1}\widehat{\boldsymbol{L}_a}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \\
&= \boldsymbol{p}_a \widehat{\boldsymbol{p}_s} \exp(-(\widehat{\boldsymbol{B}_a} + \widehat{\boldsymbol{B}_s})t) \frac{\widehat{\boldsymbol{L}_a}}{\boldsymbol{p}_a \widehat{\boldsymbol{p}_s}(\widehat{\boldsymbol{B}_a} + \widehat{\boldsymbol{B}_s})^{-1}\widehat{\boldsymbol{L}_a}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \widehat{\boldsymbol{e}_a} e_s
\end{aligned}
$$

Note that the expression in the denominator is the probability that Arrival happen before Service completion and hence is some scalar (less than 1), say α. The effect of conditioning on the fact that arrival occurs before service event, is that the Arrival processes gets effectively accelerated (from $\boldsymbol{L}_a$ to $\frac{\boldsymbol{L}_a}{\alpha}$). This in essence is the effect of knowing that additional piece of information. If we consider this as a new matrix exponential process, we no longer have the usual equality $\boldsymbol{Be} = \boldsymbol{Le}$ since $(\widehat{\boldsymbol{B}_a} + \widehat{\boldsymbol{B}_s})\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \neq \frac{\boldsymbol{L}_a}{\alpha}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$. But nonetheless this is a valid matrix exponential density. It can easily be seen that the integral of the above conditional density from 0 to ∞ equals 1.

III. ME Representation for the Length of a Sample Path

Consider a sample path during a busy period where immediately after the start of a busy period, we have an arrival followed by a departure event. The length of this sample path is the convolution of two stochastic processes, representing the occurence of an arrival event followed by a departure event.

Pr[Arrival followed by Departure = t] =

$$
\int_{t_1=0}^{t} \boldsymbol{p}_{bp} \exp(-(\widehat{\boldsymbol{B}_a + \boldsymbol{B}}_s)t_1)\frac{\boldsymbol{L}_a}{\alpha_1} \exp(-(\widehat{\boldsymbol{B}_a + \boldsymbol{B}}_s)(t - t_1))\frac{\boldsymbol{L}_s}{\alpha_2} dt_1 \tag{1}
$$

where $\alpha_1 = \boldsymbol{p}_{bp}(\widehat{\boldsymbol{B}_a + \boldsymbol{B}}_s)^{-1}\widehat{\boldsymbol{L}_a}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s = \boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$ and $\alpha_2 = \frac{\boldsymbol{p}_{bp}\boldsymbol{H}_a}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}\boldsymbol{H}_s\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$

We show in this section that that above density of a partial path can be written in matrix exponential notation using the following $< \boldsymbol{p}_{pp}, \boldsymbol{B}_{pp}, \boldsymbol{L}_{pp}, \boldsymbol{e}_{pp} >$ where,

$\boldsymbol{p}_{pp} = \left[\begin{array}{cc} \boldsymbol{p}_{bp} & 0 \end{array}\right]$

$$\boldsymbol{B}_{pp} = \left[\begin{array}{cc} (\widehat{\boldsymbol{B}_a + \boldsymbol{B}_s}) & \frac{-\boldsymbol{L}_a}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \\ 0 & (\widehat{\boldsymbol{B}_a + \boldsymbol{B}_s}) \end{array}\right], \boldsymbol{L}_{pp} = \left[\begin{array}{cc} 0 & 0 \\ 0 & \frac{\boldsymbol{L}_s}{\frac{\boldsymbol{p}_{bp}\boldsymbol{H}_a}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}\boldsymbol{H}_s\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \end{array}\right], \boldsymbol{e}_{pp} = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \\ \widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{array}\right]$$

Let us look at the Laplace transforms of the above two representations. The Laplace transform of the first representation, the convolution of two processes, can be written as the product of the individual transforms. Also note that the two processes are dependent on each other and the dependency is carried using the matrix $\boldsymbol{L}_a$ and as such the joint laplace transform can be written as

$$F^*(s)_1 = \boldsymbol{p}_{bp}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_s}{\alpha_2}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$$

The Laplace transform using the second representation for our density is given by,

$$F^*(s)_2 = \boldsymbol{p}_{pp}(\boldsymbol{B}_{pp} + \boldsymbol{sI})^{-1}\boldsymbol{L}_{pp}\boldsymbol{e}_{pp}$$

Using the fact that the inverse of a block matrix can be written as,

$$\left[\begin{array}{cc} A & B \\ C & D \end{array}\right]^{-1} = \left[\begin{array}{cc} I & -A^{-1}B \\ 0 & I \end{array}\right]\left[\begin{array}{cc} A^{-1} & 0 \\ 0 & S_A^{-1} \end{array}\right]\left[\begin{array}{cc} I & 0 \\ -CA^{-1} & I \end{array}\right]$$

where S_A, the $Schur$ complement of A is given by $S_A = D - CA^{-1}B$

$$\begin{aligned} (\boldsymbol{B}_{pp} + \boldsymbol{sI})^{-1} &= \left[\begin{array}{cc} I & (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1} \\ 0 & I \end{array}\right]\left[\begin{array}{cc} (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} & 0 \\ 0 & (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} \end{array}\right]\left[\begin{array}{cc} I & 0 \\ 0 & I \end{array}\right] \\ &= \left[\begin{array}{cc} (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} & (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} \\ 0 & (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} \end{array}\right] \end{aligned}$$

Hence,

$$\begin{aligned} F^*(s)_2 &= \left[\begin{array}{cc} \boldsymbol{p}_{bp}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} & \boldsymbol{p}_{bp}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1} \end{array}\right]\left[\begin{array}{cc} 0 & 0 \\ 0 & \frac{\boldsymbol{L}_s}{\alpha_2} \end{array}\right]\left[\begin{array}{c} 1 \\ \vdots \\ 1 \\ \widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{array}\right] \\ &= \left[\begin{array}{cc} 0 & \boldsymbol{p}_{bp}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_s}{\alpha_2} \end{array}\right]\left[\begin{array}{c} 1 \\ \vdots \\ 1 \\ \widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{array}\right] \\ &= \boldsymbol{p}_{bp}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_a}{\alpha_1}(\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\frac{\boldsymbol{L}_s}{\alpha_2}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s = F^*(s)_1 \end{aligned}$$

Hence both the representations are equivalent, and the shown matrix exponential representation corresponds to the convolved sample path density.

IV. Weighted Laplace transform of a sample path during a Busy Period

Let us consider all possible paths during which exactly two customers are served during a busy period. As the busy period starts with the first arriving customer, there is only one such path possible, another arrival followed by two consecutive departures, i.e., "A-D-D". The probability of this path being taken is $p_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$. The matrix exponential representation for the length of this busy period is given by $<\boldsymbol{p}_{pp}, \boldsymbol{B}_{pp}, \boldsymbol{L}_{pp}, \boldsymbol{e}_{pp}>$ where,

$$\boldsymbol{p}_{pp} = \begin{bmatrix} \boldsymbol{p}_{bp} & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{B}_{pp} = \begin{bmatrix} (\widehat{\boldsymbol{B}}_a + \widehat{\boldsymbol{B}}_s) & \dfrac{-\widehat{\boldsymbol{L}_a}}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} & 0 \\ 0 & (\widehat{\boldsymbol{B}}_a + \widehat{\boldsymbol{B}}_s) & \dfrac{-\widehat{\boldsymbol{L}_s}}{\dfrac{\boldsymbol{p}_{bp}\boldsymbol{H}_a}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}\boldsymbol{H}_s\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \\ 0 & 0 & (\widehat{\boldsymbol{B}}_a + \widehat{\boldsymbol{B}}_s) \end{bmatrix}$$

$$\boldsymbol{L}_{pp} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \dfrac{\widehat{\boldsymbol{L}_s}}{\dfrac{\boldsymbol{p}_{bp}\boldsymbol{H}_a\boldsymbol{H}_s}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\boldsymbol{H}_s\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}\boldsymbol{H}_s\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s} \end{bmatrix}, \; \boldsymbol{e}_{pp} = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{bmatrix}$$

The Laplace transform for this path is given by

$$F^*(s) = \quad \boldsymbol{p}_{bp}\frac{\boldsymbol{F}_a}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}\frac{\boldsymbol{F}_d}{\dfrac{\boldsymbol{p}_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}}\frac{\boldsymbol{F}_d}{\dfrac{\boldsymbol{p}_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}{\boldsymbol{p}_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s}}\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$$

where $\boldsymbol{F}_a = (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\boldsymbol{L}_a$ and $\boldsymbol{F}_d = (\boldsymbol{B}_a \widehat{+ \boldsymbol{B}_s} + \boldsymbol{sI})^{-1}\boldsymbol{L}_s$

The weighted laplace transform for this path, conditioned by the probability of this path being taken ($p_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s$), is given by,

$$\begin{aligned} F^*(s) \text{ weighted} &= F^*(s) * p_{bp}\boldsymbol{H}_a\boldsymbol{H}_d\boldsymbol{H}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \\ &= p_{bp}\boldsymbol{F}_a\boldsymbol{F}_d\boldsymbol{F}_d\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{aligned}$$

V. Mean length of a Busy Period

Noting the similarity between the above formulation for the weighted laplace transform and the probability of a certain path being taken during a busy period, we can derive the joint transform equation for the number of customers served during a busy period and its length. We have,

$$\begin{aligned} \boldsymbol{F}(s)_0 &= \mathrm{I} \\ \boldsymbol{F}(s)_1 &= \boldsymbol{F}_a\boldsymbol{F}(s)_0\boldsymbol{F}_d\boldsymbol{F}(s)_0 \\ \boldsymbol{F}(s)_2 &= \boldsymbol{F}_a\boldsymbol{F}(s)_1\boldsymbol{F}_d\boldsymbol{F}(s)_0 + \boldsymbol{F}_a\boldsymbol{F}(s)_0\boldsymbol{F}_d\boldsymbol{F}(s)_1 \\ &\vdots \\ \boldsymbol{F}(s)_n &= \boldsymbol{F}_a[\boldsymbol{F}(s)_{n-1}\boldsymbol{F}_d\boldsymbol{F}(s)_0 + \boldsymbol{F}(s)_{n-2}\boldsymbol{F}_d\boldsymbol{F}(s)_1 + \ldots + \boldsymbol{F}(s)_0\boldsymbol{F}_d\boldsymbol{F}(s)_{n-1}] \end{aligned}$$

Z-transform of the above set of equations gives the two-dimensional transform for number served during the busy period and the length of the busy period

$$\boldsymbol{F}(s,z) = z\,\boldsymbol{I} + \boldsymbol{F}_a(s)\boldsymbol{F}(s,z)\boldsymbol{F}_d(s)\boldsymbol{F}(s,z)$$

Evaluating the joint transform at $z = 1$ and including the final departure gives the Laplace transform for the busy period duration as

$$\boldsymbol{F}(s)\boldsymbol{F}_d(s) = \boldsymbol{F}_d(s) + \boldsymbol{F}_a(s)\boldsymbol{F}(s)\boldsymbol{F}_d(s)\boldsymbol{F}(s)\boldsymbol{F}_d(s) \tag{2}$$

Substituting $\boldsymbol{F}(s)\boldsymbol{F}_d(s) = \boldsymbol{F}_T(s)$ and $\boldsymbol{F}(s)\boldsymbol{F}_d(s)\Big|_{s=0} = \boldsymbol{F}_T$, taking the derivative with respect to s and evaluating at $s = 0$, we get,

$$(\boldsymbol{F}_T(s))'\Big|_{s=0} = \left(\boldsymbol{F}_d(s)' + \boldsymbol{F}_a(s)'\boldsymbol{F}_T(s)^2 + \boldsymbol{F}_a(s)\boldsymbol{F}_T(s)'\boldsymbol{F}_T(s) + \boldsymbol{F}_a(s)\boldsymbol{F}_T(s)\boldsymbol{F}_T(s)'\right)\Big|_{s=0}$$

Using $\boldsymbol{F}_a(s)\Big|_{s=0} = \boldsymbol{H}_a$, $\quad \boldsymbol{F}_d(s)\Big|_{s=0} = \boldsymbol{H}_d$, $\quad \boldsymbol{F}_a(s)'\Big|_{s=0} = -\boldsymbol{D}\boldsymbol{H}_a$ and $\boldsymbol{F}_d(s)'\Big|_{s=0} = -\boldsymbol{D}\boldsymbol{H}_d$.

$\boldsymbol{F}'_T$ is obtained by iteration on

$$\boldsymbol{F}'_T = -\boldsymbol{D}\boldsymbol{H}_d - \boldsymbol{D}\boldsymbol{H}_a\boldsymbol{F}_T^2 + \boldsymbol{H}_a\boldsymbol{F}'_T\boldsymbol{F}_T + \boldsymbol{H}_a\boldsymbol{F}_T\boldsymbol{F}'_T$$

where, $D = (\boldsymbol{B}_a + \boldsymbol{B}_s)^{-1}$ and $\boldsymbol{F}_T = \boldsymbol{Y}\boldsymbol{H}_d$

Let τ_b represent the r.v for the length of a busy period, then

$$\begin{aligned} \mathrm{E}[\tau_b] &= -\frac{d}{ds}\left(\boldsymbol{p}_{bp}\boldsymbol{F}(s)\boldsymbol{F}_d(s)\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s\right)\Big|_{s=0} \\ &= -\frac{d}{ds}\left(\boldsymbol{p}_{bp}\boldsymbol{F}_T(s)\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s\right)\Big|_{s=0} = -\boldsymbol{p}_{bp}\boldsymbol{F}'_T\widehat{\boldsymbol{e}_a}\boldsymbol{e}_s \end{aligned}$$

VI. Simplifications in an M/M/1 case

A. Length of Busy Period

In this case

$$\boldsymbol{F}_a = \frac{\lambda}{\lambda + \mu + s} \quad \text{and} \quad \boldsymbol{F}_d = \frac{\mu}{\lambda + \mu + s}$$

Hence the Laplace transform for the length of busy period Eq.[2], simplifies to

$$\boldsymbol{F}(s) = 1 + \frac{\lambda}{\lambda + \mu + s}\boldsymbol{F}(s)\frac{\mu}{\lambda + \mu + s}\boldsymbol{F}(s)$$

Therefore,

$$\lambda\mu\boldsymbol{F}(s)^2 - (\lambda + \mu + s)^2\boldsymbol{F}(s) - (\lambda + \mu + s)^2 = 0$$

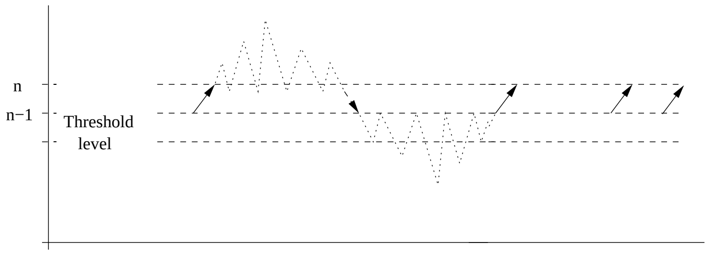
Solving for $\boldsymbol{F}\ (s)$ and selecting the appropriate root using the condition that $\boldsymbol{F}\ (s)\Big|_{s=0} = 1$ and post-multiplying with $\boldsymbol{F}_d$ gives the well know transform for busy period as,

$$\text{Transform(Busy Period)} = \frac{(\lambda + \mu + s) - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}$$

-A known result.

VII. MEAN FIRST PASSAGE TIME FOR DIFFERENT THRESHOLD LEVELS

Consider a system where the system just transitioned from level $n-1$ to level n and we are interested in the mean first passage time to reach back to level $n-1$. In this section we show the effect of correlations in arrival and service processes on the mean first passage time to go from back to this threshold level $n-1$ for different threshold levels.



This first passage time differs from a normal busy period only in the way the process starts. Once that starting vector for this "Elevated Busy Period" is known, then the rest of the analysis is similar to a normal busy period. To compute this starting vector, we consider all possible paths that result in such a transition and then compute the invariance vector. If we let $P_{n-1,n}$ denote the probability matrix that represents all the possible paths that lead the queue from level $n-1$ to level n for the first time, using a common first passage argument [9], we can compute $P_{n-1,n}$ using the following set of recurrence relations.

$$
\begin{aligned}
\boldsymbol{P}_{0,1} &= \widehat{\boldsymbol{V}_a}\widehat{\boldsymbol{L}_a} \\
\boldsymbol{P}_{1,2} &= (\boldsymbol{I} - \boldsymbol{H}_d\boldsymbol{P}_{0,1})^{-1}\boldsymbol{H}_a \\
&\vdots \\
\boldsymbol{P}_{n-1,n} &= (\boldsymbol{I} - \boldsymbol{H}_d\boldsymbol{P}_{n-2,n-1})^{-1}\boldsymbol{H}_a
\end{aligned}
$$

If we cross a given threshold $(n-1)$ and start the process in the n^{th} state with a starting vector $p_{n-1,n}$, then $\boldsymbol{Y}\boldsymbol{H}_d\boldsymbol{P}_{n-1,n}$ represents all the possible ways in which we will cross the same threshold for the first time, after dropping below the threshold. Hence the required starting vector for crossing a threshold of $n-1$ reaching n is computed as the invariance vector for the matrix $\boldsymbol{Y}\boldsymbol{H}_d\boldsymbol{P}_{n-1,n}$.

To show the effect of increase in threshold level on this first passage time, and to study the effect of correlation in the arrival process and service process on the mean of this first passage time, we use the general setup for the *MEP/MEP/1* system. For a c^2 of 25 for the arrival process and 9 for the service process, at a Utilization of 0.75 and correlation decay parameter of 0.7 (where applicable), we plot this mean first passage time as a function of the threshold level in Fig.**??**.

When both the arrival and service processes have a correlation decay parameter of 0.7, the starting vector for the correlated G/G/1 case for a transition from state 1 to state 0 is $p_{0,1} = (0.853, 0.016, 0.125, 0.005)$ and the mean length of the first passage time from level 1 to level 0 is 3.92. As we increase the threshold level, the mean first passage time from level n to $n-1$ increases and converges. In this example, the starting vector converges to $p_{n,n-1} = (0.45, 0.535, 0.006, 0.008)$ and the mean first passage time converges to 36.46 which is quite higher than the mean first passage time from level 1 to level 0. This increase in mean busy period as threshold level increases can be understood by noticing that for the queue to cross the higher threshold the internal states of the Arrival and Service processes should already be such that either the Arrival process is in its faster state or the Service process

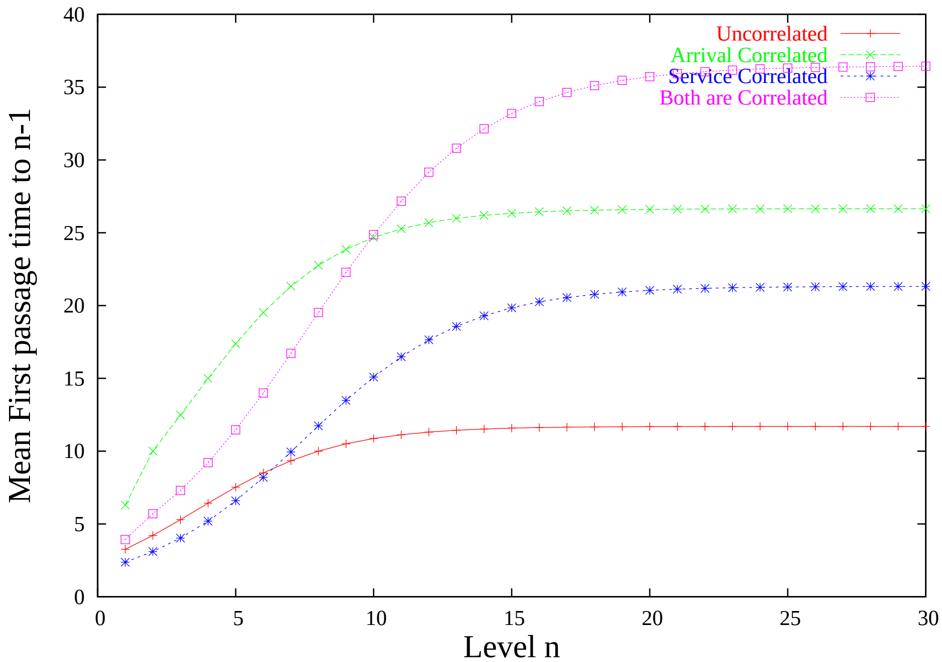
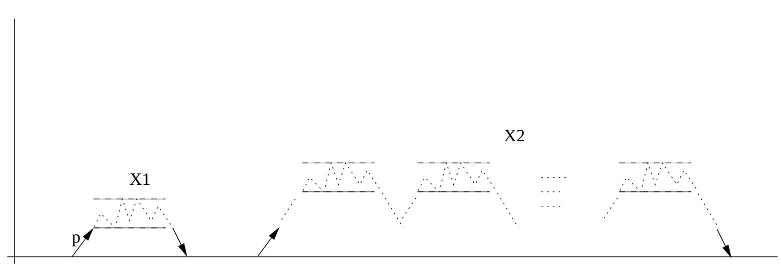


Fig. 1. Mean of Elevated Busy Periods

is in it slower state or both; and due to the correlation in these processes, the arrival and/or the service processes tend to remain in those same states for a while, which means that the transient queue length at a higher threshold is bound to increase more than in the case when the the threshold was lower. After a certain height the mean busy period converges because once the queue reaches a certain height, the probability of the Arrival process being in the slower state or the Service process being in the faster state is so low that a further increase in threshold does not effect the starting phases for the arrival and service processes.

VIII. Paths that Cross a given Level during a Busy Period

In this section we compute the probabilities of going above a height h during the first passage from level n to level $n-1$. We then show the effect of correlations in the arrival and service processes and the effect of the starting level (n), on these probabilities.



Let X_h represent all the possible paths that start at a given level and end at the same level, never going below that level and are of height atmost h. We now have the following set of recurrence relations for X_h's

$$\boldsymbol{X}_0 = \boldsymbol{I}$$
$$\boldsymbol{X}_1 = (\boldsymbol{I} - \boldsymbol{H}_a \boldsymbol{X}_0 \boldsymbol{H}_d)^{-1}$$
$$\boldsymbol{X}_2 = (\boldsymbol{I} - \boldsymbol{H}_a \boldsymbol{X}_1 \boldsymbol{H}_d)^{-1}$$

$$\vdots \quad = \quad \vdots$$

Let M_n be the r.v for the density for Maximum height during a first passage from level n to level $n-1$. Hence,

$$\text{Prob}[M_n \leq h\ ] = \boldsymbol{p}\boldsymbol{X}_{h-1}\boldsymbol{H}_d\boldsymbol{e} \qquad\qquad h \geq 1$$

The probability that during a busy period a given level is crossed is hence computed as

$$\text{Prob}[M_n > h\ ] = 1 - \boldsymbol{p}\boldsymbol{X}_{h-1}\boldsymbol{H}_d\boldsymbol{e} \qquad\qquad h \geq 1$$

The starting vector $\boldsymbol{p}$ for starting at different levels is computed using the approach presented in the previous section. The table below shows the effect of both the correlations in the arrival and service processes as well as the effect of changing the base level on these probabilities. The correlation decay parameter gamma is set to 0.7 and a squared coefficient of variation of 9 is used for both the Arrival and Service processes. Utilization is set to 0.75.

	MM1	MM1	GG1	GG1
	From 1-0	From 10-9	From 1-0	From 10-9
h	Prob. that Height greater than h			
1	0.428571	0.428571	0.369322	0.620431799
2	0.243243	0.243243	0.212039	0.501704526
3	0.154286	0.154286	0.140600	0.442551601
4	0.103713	0.103713	0.101558	0.406558265
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	0.014699	0.014699	0.039565	0.316721063
Mean length of the busy period?				
Mean length	0.799	0.799	2.07	16.32
Mean nos. served in a busy period?				
Mean nos.	4	4	10.38	70.81
csquare	5.25	5.25	61.2	10.8

Some very interesting numbers can be seen in the table above. For example, the probability that the queue grows above a height of 10 during a busy period for a GG1 system for a transition from level 1 to level 0 is 0.03956 where as the same probability for a transition from level 10 to level 9 is 0.3167 which is orders of magnitude higher. The effect of this can be clearly seen in both the mean length of the busy period which increases from 2.07 to 16.32 and the mean number served during a busy period which increases from 10.38 to 70.81. This is purely the effect of the increase in base level. Also note the increase in csquare for the the mean number served in a busy period increases from 5.25 in an MM1 case to 61.2 in a GG1 case; this is the sum effect of the correlations and variances in the arrival and service processes. Also there are some not so obvious numbers such as the decrease in csquare of a GG1 system as the base level changes from 1 to 10; but perhaps the reason for this is that at a higher level there is a lower chance of having fewer number of customers served, i.e., the mean number served is high (70.81) with relatively small variance of 10.8 compared to a high variance of 61.2 in the case where the base level is 1.

References

[1] Abate, J. and Whitt, W. (1992). *Numerical Inversion of Probability Generating Functions*. Operations Research Letters, Vol. 12, :4, 245-251

[2] Choudhury, G. L. Lucantoni, D. M.and Whitt, W. (1994). *Multidimensional transform inversion with applications to the transient M/G/1 queue*. Ann. Appl. Prob. 4, 719– 740.

[3] L.-M. Le Ny, B. Sericola. *Busy Period Distribution of the BMAP/PH/1 Queue*. Proceedings of the 9th International Conference on Analytical and Stochastic Modelling Techniques (ASMT), Darmstadt, Germany, June 2002.

[4] Takacs, *Introduction to the Theory of Queues* See pages 31-37.

[5] Takacs, *A Generalization of the Ballot Problem and its Applications in the Theory of Queues*. Journal of the American Statistical Association, Vol.57, No. 298 (June, 1962), Pg. 327-337.

[6] Lucantoni D, *The* $BMAP/G/1$ *Queue: A Tutorial*. Lecture Notes in Computer Science, Vol.729, 1993, Pg. 330-358.

[7] Lucantoni, David, *Further transient analysis of the $BMAP/G/1$ Queue*. Special issue in honor of Marcel F. Neuts. Comm. Statist. Stochastic Models 14 (1998), no. 1-2, 461–478.
[8] M. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopking Univ. Press, 1981.
[9] L. Lipsky, *Queueing Theory: A Linear Algebraic Approach*. New York: MacMillan, 1992.
[10] P. Peart and W.J. Woan, *Dyck paths with no peaks at height k* , Journal of Integer Sequences, vol. 4, 2001, Article 01.1.3A.
[11] R. Cooper, *Introduction to Queueing Theory*. Page 231
[12] K. Mitchell, *Constructing Correlated Sequence of Matrix Exponentials with Invariant First-Order Properties*, Operations Research Letters, vol. 28 $\#1$, pages 27-34, 2001.
[13] J. Abate and W. Whitt, *Approximations for the M/M/1 Busy Period*. Queueing Theory and its Applications, Liber Amicorum for Professor J. W. Cohen North-Holland, Amsterdam, 1988, pp. 149–191
[14] O. J. Boxma and V. Dumas, *The busy period in the fluid queue*. Centrum voor Wiskunde en Informatica (CWI), Amsterdam, Netherlands. PNA-R9718, 1997.
[15] S. Asmussen and M. Bladt, *A sample path approach to mean busy periods for Markov-modulated queues and fluids*. Adv. Appl. Prob. 26, No. 4, 1117-1121(1994).
[16] M. Agarwal, *Distribution of number served during a busy period of GI/M/1/N queues-lattice path approach*. Journal of Statistical Planning and Inference 101, 2002, Pg. 7–21.