Stabilizing Performance in a Single-Server Queue with Time-Varying Arrival Rate: APPENDIX

Ward Whitt

Department of Industrial Engineering and Operations Research,
Columbia University, New York, NY, 10027 {ww2040@columbia.edu}

February 2, 2015

Abstract

We consider a class of general $G_t/G_t/1$ single-server queues, including the $M_t/M_t/1$ queue, with unlimited waiting space, service in order of arrival and a time-varying arrival rate, where the the service rate at each time is subject to control. We study the rate-matching control, where the the service rate is made proportional to the arrival rate. We show that the model with the rate-matching control can be regarded as a deterministic time transformation of a stationary $G/G/1$ model, so that the queue length distribution is stabilized as time evolves. However, the time-varying virtual waiting time is not stabilized. We show that the time-varying expected virtual waiting time with the rate-matching service-rate control becomes inversely proportional to the arrival rate in a heavy-traffic limit. We also show that no control that stabilizes the queue length asymptotically in heavy-traffic can also stabilize the virtual waiting time. Then we consider two square-root service-rate controls and show that one of these stabilizes the waiting time when the arrival rate changes very slowly relative to the average service time, so that a pointwise stationary approximation is appropriate.

Keywords: stabilizing performance, queues with time-varying arrival rates, nonstationary queues, heavy-traffic limits, single-server queues with time-varying arrival rates, service-rate controls, heavy-traffic scaling, pointwise stationary approximations.

1 Introduction

This is an appendix to the main paper [6]. In §2 we present simulation experiments for $G_t/G_t/1$ models where the arrival and service processes are non-Poisson, i.e., not M_t. In §3 we present a proof of Lemma 1 of [3] using the argument to prove Theorem 7.1 of [6]. In §4 we present the results of additional simulation experiments.

2 Simulation Experiments for Non-Markovian $G_t/G_t/1$ Models

We also conducted simulation experiments for non-Markovian $G_t/G_t/1$ models, using the same sinusoidal arrival rate functions as for the main $M_t/M_t/1$ models. For related work on conventional staffing (choosing the number of servers) for many-server $G_t/G_t/s_t$ models having non-M_t arrival processes constructed in the same way, see [1].

Here we consider arrival processes as defined in §3 of [6], where the base process N_a is a non-Poisson renewal process. The variability parameter c_a^2 is then the squared coefficient of variation (scv, variance divided by the square of the mean) of an interarrival time. We considered renewal processes with hyperexponential (H_2) and Erlang (E_2) interarrival times. The H_2 interarrival times are more variable than an exponential and were chosen to have $c_a^2 = 2.0 > 1$, while the E_2 interarrival times are less variable than an exponential and have $c_a^2 = 0.5 < 1$. For the H_2 distribution, the third parameter beyond the mean and scv c_a^2 was chosen to yield balanced means, as in p. 137 of [4].

Similarly we let the service requirements come from H_2 and E_2 renewal processes. Hence, we considered $G_t/G_t/1$ models where neither the arrivals nor the service is M_t.

In particular, Figures 1 and 2 show the performance of the rate-matching service-rate control in (2.1) of [6] applied to the model in which N_a is an H_2 renewal proceess, while the service requirements are H_2 and E_2, respectively. The arrival rate function is the same as for Figure 1 of [6]. Just as for the $GI/GI/1$ models, see §5.1 of [5], the stable mean value $EQ(t)$ tends to be higher for the more variable H_2 service-requirement distribution than for E_2.

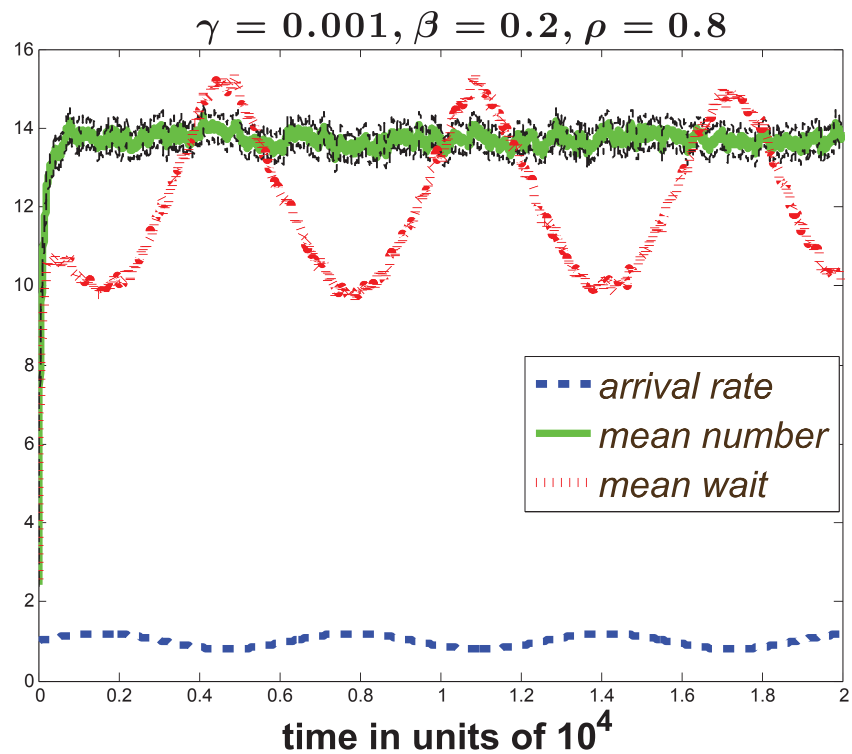

Figure 1: Simulation estimates of the time-varying mean number in the system, $E[Q(t)]$, and the mean waiting time, $E[W(t)]$, for the $G_t/G_t/1$ model with the rate-matching control in (2.1). The arrival process is constructed from a rate-1 renewal process with H_2 inter-renewal times having scv $c^2 = 2.0$. Again there is the same sinusoidal arrival rate function $\lambda(t) \equiv 1 + \beta \sin \gamma t$ with $\beta = 0.2$ and $\gamma = 0.001$. The service times are i.i.d. H_2 random variables, also with $c_s^2 = 2.0$.

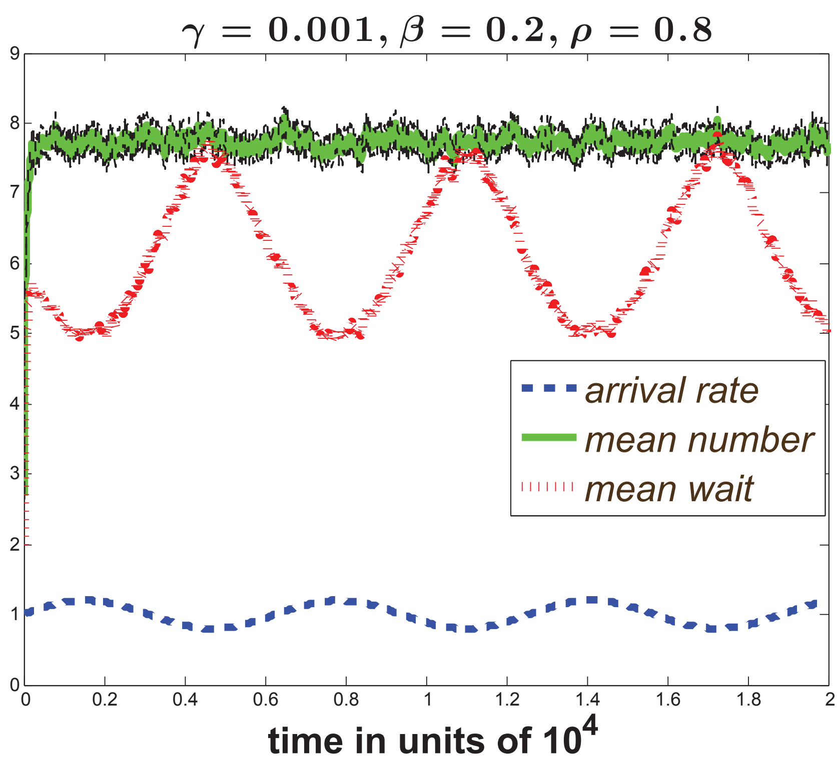

Figure 2: Simulation estimates of the time-varying mean number in the system, $E[Q(t)]$, and the mean waiting time, $E[W(t)]$, for the $G_t/G_t/1$ model with the rate-matching control in (2.1). The arrival process is constructed from a rate-1 renewal process with H_2 inter-renewal times having scv $c^2 = 2.0$. Again there is the same sinusoidal arrival rate function $\lambda(t) \equiv 1 + \beta \sin \gamma t$ with $\beta = 0.2$ and $\gamma = 0.001$. The service times are i.i.d. E_2 random variables with $c_s^2 = 0.5$.

Similarly, Figures 3 and 4 show the performance of the second square-root service-rate control in (2.4) of [6] applied to the model in which N_a is an H_2 renewal process, while the service requirements are H_2 and E_2, respectively. The arrival rate function is the same as for Figure 3 of [6], which is the same as in Figures 1 and 2 here.

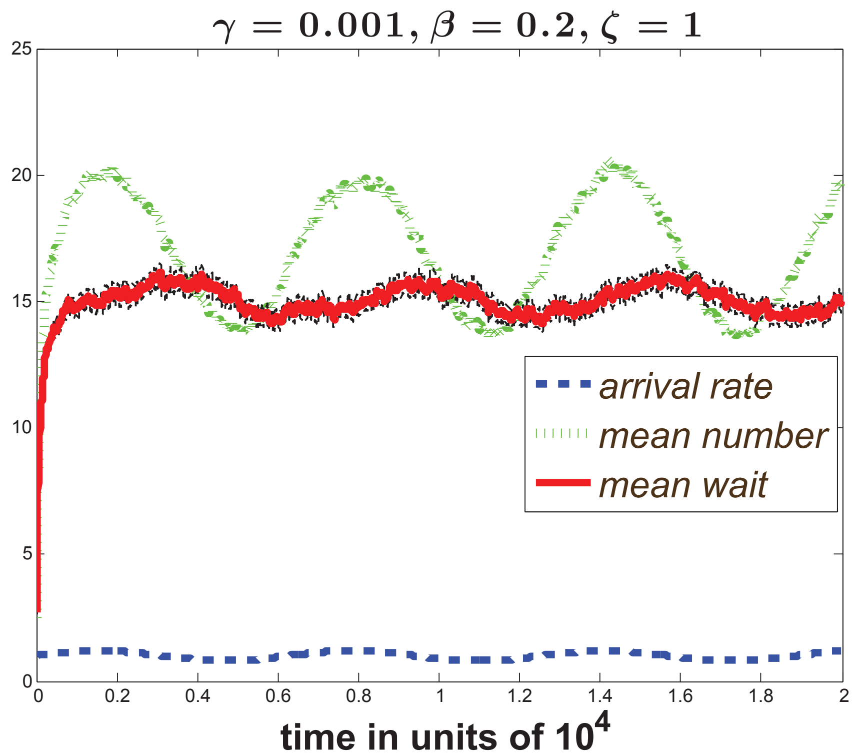

Figure 3: Simulation estimates of the time-varying mean number in the system, $E[Q(t)]$, and the mean waiting time, $E[W(t)]$, for the $G_t/G_t/1$ model with the second square-root control in (2.4). The arrival process is constructed from a rate-1 renewal process with H_2 inter-renewal times having scv $c^2 = 2$. Again there is the same sinusoidal arrival rate function $\lambda(t) \equiv 1 + \beta \sin \gamma t$ with $\beta = 0.2$ and $\gamma = 0.001$. The service times are i.i.d. H_2 random variables, also with $c_s^2 = 2.0$.

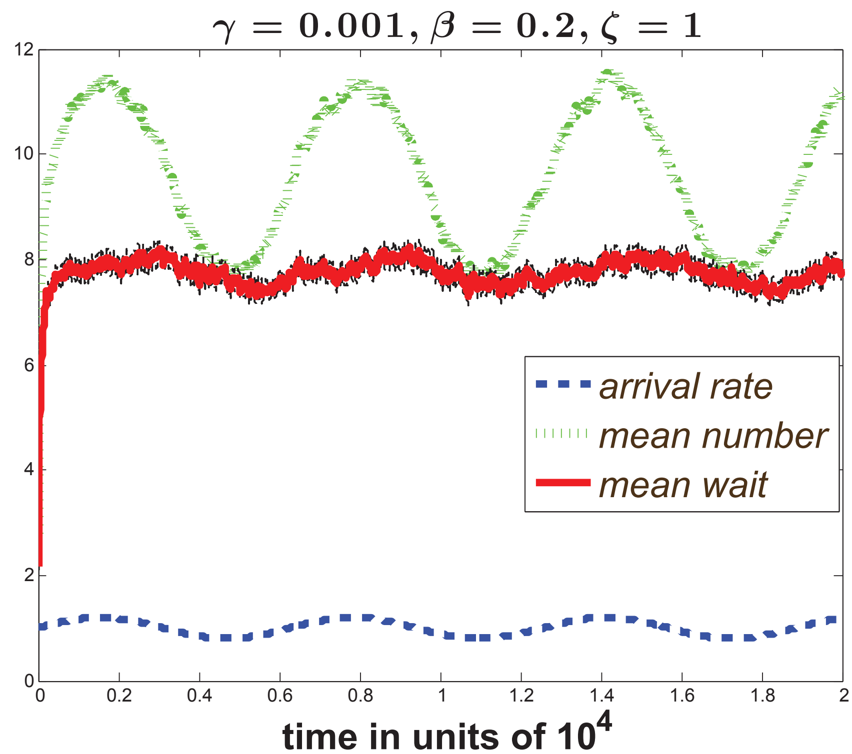

Figure 4: Simulation estimates of the time-varying mean number in the system, $E[Q(t)]$, and the mean waiting time, $E[W(t)]$, for the $G_t/G_t/1$ model with the second square-root control in (2.4). The arrival process is constructed from a rate-1 renewal process with H_2 inter-renewal times having scv $c^2 = 2$. Again there is the same sinusoidal arrival rate function $\lambda(t) \equiv 1 + \beta \sin \gamma t$ with $\beta = 0.2$ and $\gamma = 0.001$. The service times are i.i.d. E_2 random variables with $c_s^2 = 0.5$.

3 Proof of Lemma 1 of Liu and Whitt [3]

The impossibility result in Theorem 7.1 of the main paper [6] showing that it is not possible to simultaneously stabilize the mean queue length $EQ(t)$ and the mean wait $EW(t)$ is similar to Lemma 1 of [3], which was used to prove a similar asymptotic impossibility result in Corollary 1 of [3]. We now show that a minor modification of our proof of Theorem 7.1 in [6] can be used to provide a short proof of Lemma 1 of [3].

Lemma 3.1 (*Lemma 1 from [3]*) *Given that F is a cdf, $F^c(x) \equiv 1 - F(x)$ and λ is a time-varying arrival rate function with $0 < \lambda_L \leq \lambda(t)$ for all t as in Assumption 10 of [2], which is assumed in [3], if*

$$m(t) = \int_0^w \lambda(t-x) F^c(x)\, dx, \quad t \geq w, \tag{3.1}$$

is a positive constant for all $t \geq w$, then λ is a constant function.

Proof. Given that $m(t+\epsilon) = m(t)$ for any $\epsilon > 0$ and any $t \geq w$ as in (3.1), for any for any $\epsilon > 0$, we can write

$$m(t+\epsilon) - m(t) = 0 = \int_0^w [\lambda(t+\epsilon-x) - \lambda(t-x)] F^c(x)\, dx, \quad t \geq w, \tag{3.2}$$

which is equivalent to

$$\begin{aligned} 0 &= \int_0^w \lambda(t-x)[F^c(x-\epsilon) - F^c(x)]\, dx, \quad t \geq w, \\ &= \int_0^\infty \lambda(t-x)[P((A \wedge w) + \epsilon > x) - P(A \wedge w > x)]\, dx, \quad t \geq w, \end{aligned} \tag{3.3}$$

where A is a random variable with cdf F and $A \wedge w \equiv \min\{A, w\}$, but that is not possible because, by the tail integral formula for the mean and the lower bound on λ,

$$\begin{aligned} \int_0^\infty \lambda(t-x)[P((A \wedge w) + \epsilon > x) - P(A \wedge w > x)]\, dx &\geq \lambda_L(E[(A \wedge w) + \epsilon] - E[A \wedge w] \\ &= \lambda_L \epsilon > 0. \blacksquare \end{aligned} \tag{3.4}$$

4 Other Simulation Experiments

In this section we include some more simulation results for $M_t/M_t/1$ models, supplementing the main paper [6]. Figure 5 shows the performance of the rate-matching control in (2.1) of [6] for $\gamma = 1.0$, supplementing the results in Figures 1, 6 and 8 for $\gamma = 0.001$, 0.1 and 10.0, respectively.

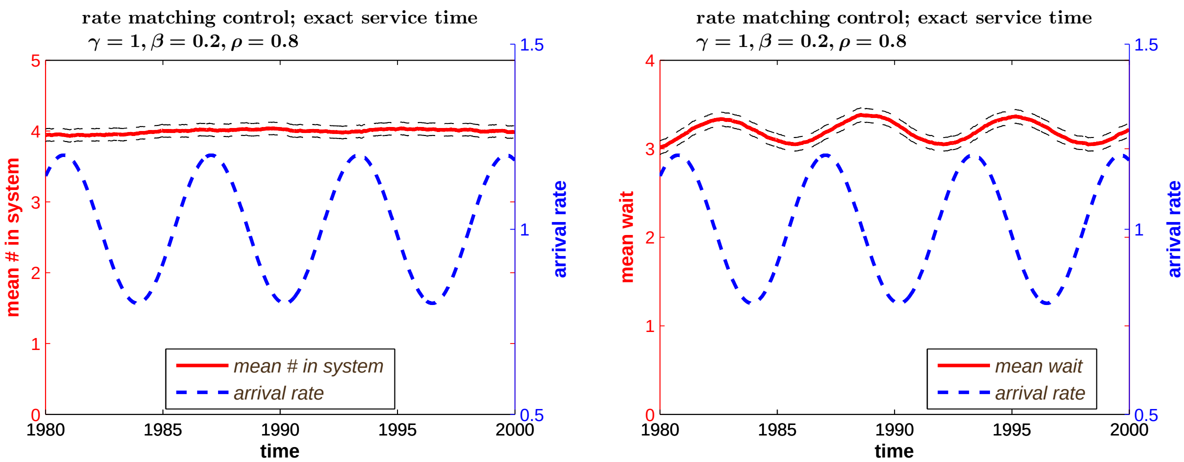

Figure 5: Simulation estimates of the time-varying mean number in the system, $E[Q(t)]$ (left), and the mean waiting time, $E[W(t)]$ (right), for the $M_t/M_t/1$ model with sinusoidal arrival rate function $\lambda(t) \equiv 1 + \beta \sin \gamma t$ with $\beta = 0.2$ and $\gamma = 1.0$ with the rate-matching control in (2.1) of [6].

Acknowledgement

The author gratefully acknowledges support from NSF grant CMMI grant 1265070 and the help of Columbia doctoral student Ni Ma, including all reported numerical and simulation results.

References

[1] B. He, Y. Liu, and W. Whitt. Staffing a service system with non-Poisson nonstationary arrivals. Columbia University, working paper, 2015.

[2] Y. Liu and W. Whitt. The $G_t/GI/s_t + GI$ many-server fluid queue. *Queueing Systems*, 71:405–444, 2012.

[3] Y. Liu and W. Whitt. Stabilizing customer abandonment in many-server queues with time-varying arrivals. *Oper. Res.*, 60(6):1551–1564, 2012.

[4] W. Whitt. Approximating a point process by a renewal process: two basic methods. *Oper. Res.*, 30: 125–147, 1982.

[5] W. Whitt. The queueing network analyzer. *Bell System Technical Journal*, 52(9):2779–2815, 1983.

[6] W. Whitt. Stabilizing performance in a single-server queue with time-varying arrival rate. Columbia University, http://www.columbia.edu/∼ww2040/allpapers.html, 2014.