REAL-TIME DELAY ESTIMATION BASED ON DELAY HISTORY SUPPLEMENTARY MATERIAL

by

Rouba Ibrahim and Ward Whitt

IEOR Department
Columbia University
{rei2101, ww2040}@columbia.edu

Abstract

Motivated by interest in making delay announcements to arriving customers who must wait in call centers and related service systems, we study the performance of alternative real-time delay estimators for the delay before entering service of an arriving customer based on recent customer delay experience. We characterize performance by the bias and the mean squared error (MSE) We do analysis and conduct simulations for the standard $GI/M/s$ multi-server queueing model, emphasizing the case of large s. The main estimators considered are: (i) the delay of the last customer to enter service (LES), (ii) the delay experienced so far by the customer at the head of the line (HOL), and (iii) the delay experienced by the customer to have arrived most recently among those who have already completed service (RCS). We compare these delay-history estimators to the estimator based on the queue length (QL), which requires knowledge of the mean interval between successive service completions in addition to the queue length. We obtain analytical results for the conditional distribution of the delay given the observed HOL delay. An approximation for the mean value of that conditional distribution, which is not the observed delay, serves as a refined estimator, yielding lower MSE than the direct estimator. We show that the MSE relative to the square of the mean is asymptotically negligible for all three candidate delay estimators in the many-server and classical heavy-traffic limiting regimes.

Keywords: delay estimation, real-time delay estimation, delay prediction, delay announcements, many-server queues, call centers, heavy traffic.

May 28, 2007

Contents

1. Introduction

This is a supplement to the main paper, Ibrahim and Whitt (2006), with the same title. Here, we report the results of simulation experiments supporting some of the theoretical results of the main paper. But, we also go beyond that and use simulation to gain insight into the performance of candidate delay estimators for which we have not obtained closed mathematical expressions.

The Main Paper. In the main paper, we present and analyze several real-time delay estimators based on recent customer delay history, in the standard $GI/M/s$ model. In the framework of the $GI/M/s$ model, we were able to obtain tractable expressions for some of our candidate delay estimators. Our research is motivated by applications to call centers but could be applied to virtually any service system where customers are subject to delays prior to receiving service, especially when they are unable to approximate the waiting time themselves (e.g., service systems with invisible queues).

We propose seven different delay estimators. A detailed description of these estimators can be found in section 2 of the main paper. In a nutshell, these estimators can be grouped into three categories. The first category holds the two useful reference estimators, QL and NI. QL (Full-Information Queue-Length Delay Estimator) is expected to perform better than all other partial-information delay estimators. NI (No-Information Steady-State Estimator) is expected to perform worse than the other estimators since they all assume some state information beyond the model. Any estimator performing consistently worse than NI is not worth serious consideration.

The second category holds LES and HOL which both announce, as estimates, delays of customers who haven't yet completed service. We therefore put them in the same category. HOL (Head-Of-The-Line Estimator) can be used as an approximation to LES (Last Customer to Enter Service). The reader interested in the validity of this approximation is referred to section 4 of the main paper. Here, we report simulation results that support this approximation.

The third category holds the LCS, RCS and RCS$-c\sqrt{s}$ estimators. These estimators all announce delays of customers who have already completed service. That is why we put them in the same category. LCS (Last Customer to Complete Service), RCS (Most Recent Arrival to Complete Service) and RCS$-c\sqrt{s}$ (Most Recent Arrival Among the Last $c\sqrt{s}$ Customers to Complete Service) are hard to analyze mathematically. Their analysis remains incomplete at this point; simulation thus plays a key role in studying the performance of these alternative estimators.

Organization of this Supplement. We start in section 2 by describing our simulation experiments. In section 3, we report simulation results comparing the performances of our delay estimators in the $GI/M/s$ model for three different interarrival time distributions. We measure the performance of a delay estimator by computing the corresponding *average squared error* (ASE). In section 4, we go beyond measuring the overall performance of a given estimator and report average square errors conditional on the level of actual delay observed. By actual delay, we mean the measured delay of a customer who has been given a delay estimate upon arrival. The computation of these conditional errors is detailed in that section. There, we see interesting results that weren't captured in the overall comparisons of section 3. In section 5, we focus on the efficiency of the LES estimator. In particular, we report simulation results supporting the approximation for the MSE of the LES estimator, which can be found in section 4 of the main paper. In section 6, we report simulation results that support another theoretical result of the main paper: the relative efficiency of HOL as compared to QL. This result can be found in section 4 of the main paper. In section 7, we study the effect of using the delay history of customers who have completed service. That is, we vary the amount of information used and make observations. We consider a family of delay estimators RCS$-f(s)$ for different functions f of the number of servers s and compare these estimators' efficiencies. Our aim here is to show that, consistent with heavy-traffic analysis, it is enough to use the delay history of the last $c\sqrt{s}$ customers, for some constant c which we determine by simulation. In section 8, we look more carefully at the distribution of our delay estimators, studying in particular the distribution of $W_{HOL}(w)$ and testing approximations in the main paper. In section 9, we draw conclusions. In section 10, we produce all tables and figures relevant to the sections of this supplement.

2. Description of the Simulation Experiments

In this section we describe our simulation experiments. Our adopted model in the main paper is the standard $GI/M/s$ queueing model with s homogeneous servers working in parallel, an unlimited waiting room and the first-come first-served service discipline. We thus let the service times V_n be independent and identically distributed (i.i.d.) exponential random variables with mean $E[V] = 1$. Fixing the mean service time as such is done without loss of generality since the service rate can be made equal to 1 by a proper choice of time units. We let the interarrival times U_n be i.i.d. positive random variables with a non-lattice cumulative distribution function (cdf) F. We measure the variability of a distribution by computing its *squared coefficient of variation* (SCV) defined as $c_a^2 = Var[U]/E[U]^2$.

In this supplement to the main paper, we use simulation to validate some of the theoretical results of the main paper. We also use simulation to gain insight into the performance of candidate delay estimators that are hard to analyze mathematically, even within the framework of the relatively simple $GI/M/s$ model. In all of our experiments, we measure the performance of a delay estimator by computing the *average squared error* (ASE), which is defined by:

$$ASE \equiv \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2, \qquad (2.1)$$ {ASEEq}

where a_i is the actual delay observed, p_i is the estimated delay announced according to the given scheme and n is the number of customers that were given a delay announcement. Throughout, we only give delay announcements to customers that had to wait in line for a positive amount of time before receiving service. In other words, we will always have $a_i > 0$ in (2.1) above. For each delayed customer i, we record the delay estimate p_i, announced according to a certain delay estimator and the actual delay observed a_i. We then average the square of the differences between these two over all customers who were given a delay estimation. The ASE computed as such approximates the *mean squared error* (MSE) of the estimators especially when the samples are sufficiently large. This motivates our using simulation experiments to further validate the closed form expressions we obtained for the MSE of some delay estimators in the main paper.

We now describe our simulation experiments. Our simulations are *steady-state discrete-event* simulations since the performance measures we aim to calculate are steady-state performance measures. Therefore, we could potentially face two problems: (1) *Initial Bias* and (2) *Autocorrelations*. The initial bias is the difference between the expected value of the statistical estimator, the ASE, and the quantity it is estimating, the MSE. This happens because the system is not started in steady-state. A partial solution to this problem is to delete some initial segment of the data, i.e., to have a warmup period which we later discard. To assess the impact of the initial bias in our simulations, we simulated the $M/M/900$, $H_2/M/900$ and $D/M/900$ queuing models and varied the length of the warmup period reporting, in each case, the computed point estimates of the ASEs for the alternative delay estimators. We define the length of a simulation run in terms of number of events. We define an event to be either an arrival or a service completion. That is, a simulation run length of 5 million events is terminated when the sum of the number of arrivals and the number of service completions in that run is equal to 5 million. To assess the importance of the initial bias, we simulated the three models above with $\rho = 0.95$. We chose a large number of servers, s, and a large traffic intensity, ρ, because it is known that the initial bias effect is likely to be more important for large s and high ρ. We simulated our models four times with: (1) 5×10^6 events and no warmup period, (2) 5×10^6 events and a warmup period of 5×10^5 events, (3) 5×10^6 events and a warmup period of 1×10^6 events and (4) 5×10^6 events and a warmup period of 2×10^6 events. We compared the point estimates of the ASEs obtained in each case and found that these estimates do not differ by a significant amount. Indeed, not having an initial warmup period caused a bias of at most 8 percent, reported for the $H_2/M/900$ model. Since we didn't detect significant initial transient effect, our simulation replications in this supplement are done without deleting an initial portion of the data, i.e., with no warmup period.

Autocorrelations are the correlations between the successive values of the process. The effect of autocorrelations can be reduced by using the method of independent replications. We use the method of independent replications to generate point and confidence interval estimates for the ASE of an estimator in a given model (for a chosen ρ). We simulate each model (for each ρ) for 10 independent replications. The length of these runs depends on the particular model and the traffic intensity at hand. In general, when the traffic intensity ρ is high (but still, $\rho < 1$), there tend to be large fluctuations and thus high variability in which case the simulation run need be longer. Also, simulation runs need be longer as the variability in the arrival process increases, especially for larger values of s. We determine whether simulation run lengths are roughly appropriate by comparing our simulation estimates to exact analytical values we have for the model at hand. For example, in our simulations for the $M/M/s$ queue, we compared our point estimates for the ASEs of QL and NIE to the well-known exact values of the MSEs of these two estimators. As a result, we varied our simulation run lengths from 5 million events when both ρ and s are small (For example, $\rho = 0.85$, $s = 10$) to 7, 10 and 15 million as s, ρ and the variability in the arrival process increase (for example, we considered a length of 15 million for each replication of the $H_2/M/900$ model with $\rho = 0.98$). For more on the estimation of simulation run lengths in queueing simulations, see Whitt (1989). In each replication k, we return the value of

$$\overline{ASE_k} = \frac{1}{n_k}\sum_{i=1}^{n_k}(a_i - p_i)^2 \ , \qquad (2.2) \quad \{est1\}$$

where n_k is the size of our sample in replication k ($k = 1, 2, ..., 10$). We then average $\overline{ASE_k}$ over all 10 replications :

$$\widehat{ASE} = \sum_{k=1}^{10}\overline{ASE_k} \qquad (2.3) \quad \{est2\}$$

Everywhere in this supplement, we report $\widehat{ASE}$ as our point estimate to the ASE of the given estimator in that particular model (for a chosen ρ). The corresponding interval estimate for $\widehat{ASE}$ is constructed as follows: given our sample $\{\overline{ASE_k},\ k = 1, 2, ..., 10\}$, we compute the sample variance:

$$s^2 = \frac{1}{9}\sum_{k=1}^{10}(\overline{ASE_k} - \widehat{ASE})^2 \qquad (2.4) \quad \{est3\}$$

Our interval estimate is then given by:

$$\widehat{ASE} \pm \frac{s}{\sqrt{10}}t_{\alpha/2,9} \ , \qquad (2.5) \quad \{ci\}$$

where α is the chosen level of significance and 9 is the number of degrees of freedom for the t-statistic (we assume normality of the data). In our tables for the ASEs, we report 95 percent confidence intervals for our point estimates, which we constructed as explained above.

In our simulations, we vary three different parameters: (1) the number of servers s, (2) the traffic intensity $\rho = E[V]/sE[U] = 1/sE[U]$ and (3) the variability of the arrival process which we measure by the SCV of the interarrival time distribution. We are particularly interested in the case of large s since we are motivated by applying this research to call centers, but we will include smaller values of s as well. In particular, we will always consider the values: $s = 1, 10, 100, 400$ and 900. Since delay announcements are more relevant when customer delays are long, we will mostly be focusing on high traffic intensities in our simulations. In general, as we increase the number of servers, we need increase the traffic intensity further in order to see large customer delays. For example, in the extreme case of $s = 900$, we only consider traffic intensities $\rho \geq 0.93$. To study the effect of variability in the arrival process, we vary the interarrival time distribution F to consider both low variability and high variability distributions. In particular, we consider Deterministic, Poisson and H_2 arrivals with respective SCVs $0, 1$ and 4. We consider these three distributions in all of our simulations.

The simulation program was written in C. We used our C code to create Excel add-ins via XLL Plus and generated output in Excel.

3. Efficiency of the Alternative Estimators in $GI/M/s$ Models

{secUncondASE

In the main paper, we presented analytical results comparing the efficiency of some alternative delay estimators in $GI/M/s$ models. These alternative delay estimators, excepting the No-Information Steady-State estimator (NI), are random variables defined by recent customer delay experience in the model at hand. Here, we complement that analysis by presenting the corresponding empirical results. In this section, we use simulation to compare the performances of our alternative delay estimators in $GI/M/s$ models. We study the effect of variability in the arrival process by considering Poisson, Deterministic and H_2 arrivals with respective squared coefficient of variation (SCV) 1, 0 and 4. We vary the number of servers s and the load in the system, which we measure by the traffic intensity $\rho = E[V]/sE[U] = 1/sE[U]$ under our assumption that $E[V] = 1$. We are interested in studying the differences between these alternative performances. In the main paper, we made a distinction between *direct* and *refined* delay estimators. Specifically, we did this for the Last Customer to Enter Service (LES) and Head-Of-The-Line (HOL) delay estimators. The observed delay is the direct estimator, while the mean of the conditional distribution of the delay to be estimated, given the observed delay, is the refined estimator based on that same observation. For the LES and HOL delay estimators, we found that the mean values of these conditional distributions are approximately equal to ρw, when the observed delay is w, provided that w is not too small. However, since we are mostly interested in heavily loaded systems where ρ is close to 1, we expect that there won't be substantial differences in performances when refining our predictors as such. We thus only consider here the *direct* LES and HOL delay estimators.

We quantify the performance of a delay estimator by computing point and interval estimates of the *average squared error* (ASE). The ASE is defined by

$$ASE \equiv \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2 \ , \qquad (3.1)$$ {ASEEqq}

where a_i is the actual delay of the i^{th} customer, p_i is the estimated delay of that customer (the value of the delay estimator at hand) and n is the number of customers in our sample. For large samples, the ASE should agree with the *mean squared error* (MSE) in steady state. Note that we only give delay announcements to delayed customers, i.e., customers that have to wait in line prior to receiving service. Thus, we will always have $a_i > 0$ for every customer i. It is possible, however, to have $p_i = 0$ for some customer i. For example, consider the LES delay announcement. Suppose that customer i_0 joins the queue at some time t and that the last customer to have entered service prior to t could be served immediately upon her arrival. Then, $p_{i_0} = 0$ and $a_{i_0} > 0$ for our customer i_0. As the number of such customers in our sample increases, we expect the performance of our delay estimator (in this case, LES) to deteriorate. Then, we might want to adjust our delay estimator in such a way so as to avoid this discrepancy around 0. We do not deal with this issue here, however, because we are interested in heavily loaded systems where the proportion of delayed customers is high. What's more, we want to study the performances of our alternative delay estimators specifically when the actual (alternatively, announced) delays reported exceed a certain threshold. We do this in §4 by computing the conditional ASEs of the alternative delay estimators: for a given delay estimator, we compute its ASE as in (3.1) but we restrict our sample to the pairs (a_i, p_i) where a_i exceeds our threshold. That is, in §4, we will study the performances of our delay estimators conditional on the level of the waiting time. In this section, we report the value of the unconditional ASE where we include *all* the pairs (a_i, p_i) in (3.1). We generate point and confidence interval estimates for the ASEs. Let $\widehat{ASE}$ be the point estimate of the ASE that we calculate.

In addition to the ASE, we compute point estimates of the *relative ASE*(RASE). We let $\widehat{RASE}$ be the point estimate of the RASE that we calculate. Then:

$$\widehat{RASE} \equiv \widehat{ASE}/(E[W_\infty \widehat{|W_\infty} > 0])^2 \qquad (3.2)$$ {RASE}

where $[W_\infty | W_\infty > 0]$ is a random variable denoting the steady-state waiting time given that the wait is positive and $E[W_\infty \widehat{|W_\infty} > 0]$ is our computed point estimate of the mean of $[W_\infty | W_\infty > 0]$. That is, to compute point estimates for the RASE, we divide our point estimate for the ASE, $\widehat{ASE}$, by the square of our point estimate for the mean, $(E[W_\infty \widehat{|W_\infty} > 0])^2$. For large samples, the RASE should agree with the *relative mean squared error* (RMSE) in steady state. The RMSE, approximated by the RASE, is appealing because of its simple form in some cases. In particular, the RMSE of the Full-Information Queue Length Delay Estimator (QL) has a linear form in the $M/M/s$ model and the RMSE of the No-Information Steady-State estimator (NI) is equal to 1 in the $GI/M/s$ model. The simple form of the RMSE makes it appealing to use. That is why we choose to adopt it. Plotting the RMSE as a function of the traffic intensity ρ produces nice linear plots which are easy to analyze. For example, see Figures 12 and 22 below.

An alternative approach, which we do not consider here, would be to compute point estimates for the *relative root average squared error* (RRASE). If $\widehat{RRASE}$ is a point estimate of the RRASE then:

$$\widehat{RRASE} \equiv \sqrt{\widehat{ASE}}/E[\widehat{W_\infty|W_\infty} > 0] \qquad (3.3) \quad \{\texttt{why not?}\}$$

That is, a point estimate for the RRASE is given by dividing the square root of our point estimate for the ASE, $\sqrt{\widehat{ASE}}$, by our point estimate for the mean, $E[\widehat{W_\infty|W_\infty} > 0]$. For large samples, the RRASE should agree with the *root relative mean squared error* (RRMSE) in steady state. Approximating the RRMSE (by estimating the RRASE) would be useful because the MSE is in the same scale as the square of the mean, thus $\sqrt{MSE}$ is in the same scale as the mean.

We can analyze the $M/M/s$ model and get closed form expressions for the ASEs of the QL and NI delay estimators. In the $M/M/s$ model, it is well known that $[W_\infty|W_\infty > 0]$ is distributed as an exponential random variable with mean $\frac{1}{s\mu(1-\rho)} = \frac{1}{s(1-\rho)}$ when $\mu = 1$ (e.g., see section 5.14 of Cooper (1981)). For the NI estimator, the MSE (thus, the ASE) should coincide with:

$$MSE(NI) = var[W_\infty|W_\infty > 0] = \frac{1}{s^2(1-\rho)^2} \qquad (3.4) \quad \{\texttt{MSEni}\}$$

We can also analyze the performance of the QL estimator. Let $[Q_\infty|Q_\infty > 0]$ be a random variable with the conditional distribution of the steady-state queue length upon arrival given that the customer must wait before beginning service. In the $M/M/s$ model, it is known that $[Q_\infty|Q_\infty > 0] + 1$ has a geometric distribution with mean $1/(1-\rho)$. Hence,

$$E[MSE(QL)] = E[Var(W_Q(Q_\infty))] = E[Q_\infty + 1] \times \frac{1}{s^2} = \frac{1}{s^2(1-\rho)} \qquad (3.5) \quad \{\texttt{MSEql}\}$$

We expect the MSEs (hence, the ASEs) of the other estimators to fall between that of the QL estimator (best possible) and the NI (worst possible). We thus always include these two estimators as reference points for the performances of all the other estimators. In addition to this, we have closed form expressions for the RMSEs of the QL and NI delay estimators. It follows directly from the above that:

$$RMSE(QL) = E[MSE(QL)]/(E[W_\infty|W_\infty > 0])^2 = 1 - \rho\ , \qquad (3.6) \quad \{\texttt{RMSEql}\}$$

and,

$$RMSE(NI) = MSE(NI)/(E[W_\infty|W_\infty > 0])^2 = 1 \qquad (3.7) \quad \{\texttt{RMSEni}\}$$

Note here that the RMSE of the QL delay estimator is linear in ρ.

We are now ready to present our simulation results. In Tables 1-10, we display our point estimates for the ASEs and RASEs of the alternative delay estimators in the $M/M/s$ model for $s = 1, 10, 100, 400$ and 900. In Tables 11-20 and 21-30, we report these point estimates for the $D/M/s$ and $H_2/M/s$ models respectively, for $s = 1, 10, 100, 400$ and 900. We vary the traffic intensity ρ, but we focus more on heavy loads: with those loads, customer delays are longer and giving delay predictions is more meaningful. In addition to the tables, we produce figures plotting the ASE and the RASE of our alternative delay estimators as a function of the traffic intensity in the model. In Figures 1-10, we show plots for the ASE and RASE of our delay estimators as a function of the traffic intensity ρ (for a fixed number of servers s) in the $M/M/s$ model with $s = 1, 10, 100, 400$ and 900. In Figures $11-20$ and $21-30$, we do the same for the $D/M/s$ and $H_2/M/s$ models respectively.

For the $M/M/s$, $D/M/s$ and $H_2/M/s$ models considered, i.e., for all s and all ρ, we can classify the performance of our estimators as follows:

$$QL > LES \approx HOL > RCS \approx RCS - \sqrt{s} > LCS > NI$$

where ">" is to be read as "performs better than" and "≈" is to be read as "performs nearly the same but slightly better than". This order of performances holds for all models here, in the range of traffic intensities considered.

We first consider the performances of our reference estimators: QL and NI. As expected, QL always takes the lead among all estimators and NI always falls behind. In the $M/M/s$ model, we see that ASE(QL), ASE(NI), RASE(QL) and RASE(NI) agree closely with the theoretical values given in (3.5), (3.4), (3.6) and (3.7). Consequently, in the $M/M/s$ model, the ratio ASE(NI)/ASE(QL) agrees closely with $1/(1-\rho)$, as expected. Note that this ratio increases as ρ increases. This explains the deterioration in the performance of NI as the load increases. Throughout, our point estimate for RASE(NI) ≈ 1, except for the $H_2/M/900$ model under heavy loading ($\rho = 0.95$ or $\rho = 0.98$) where we see that RASE(NI) ≈ 0.9, which means that the run length evidently is not long enough. We will see in §4 that NI's performance is good compared to that of the other delay estimators when we restrict our attention to small delays only. This performance deteriorates as we consider longer delays.

We see that LES and HOL outperform the rest of the estimators (except QL). Their performances are always very close. Also,

$$ASE(HOL)/ASE(QL) \approx ASE(LES)/ASE(QL) \approx (c_a^2+1)/\rho \quad (3.8) \quad \{\texttt{approx1}\}$$

with the approximation becoming more accurate as the observed delay increase, especially for large s. This coincides with theoretical results of section 4 of the main paper, where the approximation holds as $sw \to \infty$ where w is the observed delay. We will discuss this at length in section 7 of this supplement. We note here that this explains why the performances of LES, HOL and QL are so close in the $D/M/s$ model. In this model, $c_a^2 = 0$ and the approximation becomes: $ASE(HOL)/ASE(QL) \approx ASE(LES)/ASE(QL) \approx 1/\rho \downarrow 1$ as $\rho \uparrow 1$.

The analysis in sections 4 and 5 of the main paper leads to the approximation:

$$E[MSE(LES)] \approx E[MSE(HOL)] \approx \frac{1}{s^2}\left[(c_a^2+1)^2 + \frac{((2\rho-1)c_a^2+4\rho-3)(c_a^2+1)}{2(1-\rho)} + K\right] \quad (3.9) \quad \{\texttt{approx2}\}$$

where $K = 1.5c_a^4 + 4c_a^2 + 4.5 - (2/3)v_a^3$ and $v_a^3 = E[U^3]/(E[U])^3$.
Our simulation results are roughly consistent with this approximation, which is the most accurate in the $M/M/s$ model. We leave a detailed study of the accuracy of this approximation to section 4 of this supplement.

We now consider the three delay estimators based on the delay information of customers who have completed service, namely: LCS, RCS and $RCS-\sqrt{s}$. We will study the difference in performances between RCS and $RCS-\sqrt{s}$ and the effect of delay information at length in section 8 of this supplement. For $s = 1$, these three estimators perfectly coincide. This is so because with exactly one server, the order of customer departure from the system is the same as that of customer arrival to the system. Hence, the last customer that has completed service is the most recent arrival among all customers that have completed service. On the other hand, $\sqrt{s} = 1$ implies that $RCS-\sqrt{s}$ and LCS are the same. Even when the number of servers is small enough (e.g., $s = 10$) we see that the performances of LCS, RCS and $RCS-\sqrt{s}$ are still close, see for example Tables 5, 13 and 23 for the respective results of the $M/M/10$, $D/M/10$ and $H_2/M/10$ models. But, as the number of servers increases, LCS clearly falls behind; see for example Tables 9, 19 and 29 for the $M/M/900$, $D/M/900$ and $H_2/M/900$ models. We emphasize here that customers need not depart in the order of their arrival. Indeed, with exponential service times, each of the s servers is equally likely to generate the next service completion. Thus, it is possible that the last customer to have completed service has experienced his waiting time long ago. We therefore expect the delay information of the *most recent arrival* among those customers that have completed service to be more relevant to our current customer (the system changes less over shorter periods of time). Also, when the number of servers is small we see that the performances of LES and LCS are quite close (see for example Tables 1, 11 and 21). The difference between the two becomes more pronounced as the number of servers increases and is dramatic when $s = 400$ or $s = 900$. This is so because, with exponential service times, the times between consecutive departures from service are i.i.d exponential with mean 1/s. Then as the number of servers increases, the times between consecutive departures are significantly smaller than the consecutive service times (recall our assumption of mean service time $E[V] = 1$). That is, the last customer to have completed service may have experienced his waiting time much before the last customer to enter service. We have thus explained the poor performance of LCS compared to LES, which is more pronounced as s increases.

4. Efficiency of the Estimators Conditional on the Level of Delay Observed

{secCondASE}

In the main paper, we presented analytical results comparing the efficiency of some alternative delay estimators in $GI/M/s$ models. These alternative delay estimators, excepting the No-Information Steady-State estimator (NI), are random variables defined by recent customer delay experience in the model at hand. In section 3, we used simulation to compare the performances of our alternative delay estimators in $GI/M/s$ models. There, we quantified the performance of a delay estimator by computing point estimates and confidence intervals of the *average square error* (ASE). The ASE is defined by:

$$ASE \equiv \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2 \ , \qquad (4.1)$$ {ASE}

where a_i is the actual delay of the i^{th} customer, p_i is the estimated delay of that customer (the value of the delay estimator at hand) and n is the number of customers in our sample. For large samples, the ASE should agree with the *mean squared error* (MSE) in steady state. Here, we go beyond measuring the overall performance of a given estimator and report average square errors conditional on the level of actual delay observed (a_i in equation (6.3)). We study the performances of our alternative delay estimators specifically when the actual customer delays reported are in pre-specified intervals. We do this here by computing the conditional ASEs of the alternative delay estimators: for a given delay estimator, we compute its ASE as in (6.3) but we restrict our sample to the pairs (a_i, p_i) where a_i is in a pre-specified interval. That is, we study the performances of our delay estimators conditional on the level of the waiting time. We generate point and confidence interval estimates for these conditional ASEs.

Announcing delay estimates is more relevant when the observed delays in the system are long. It's not clear what one should consider to be a long delay. Here, we compare the actual delays experienced by waiting customers to our point estimate of the mean waiting time in the system, conditional on the wait being positive. Let $E[\widehat{W|W>0}]$ be this point estimate. Here, $[W|W>0]$ is a random variable with the distribution of the steady state waiting time conditional on the wait being positive. We describe an actual delay a_i as long if $a_i > E[\widehat{W|W>0}]$. Furthermore, we study different levels of long actual delays, grouping them into 4 intervals: $(E[\widehat{W|W>0}], 2E[\widehat{W|W>0}])$, $(2E[\widehat{W|W>0}], 4E[\widehat{W|W>0}])$, $(4E[\widehat{W|W>0}], 6E[\widehat{W|W>0}])$ and $> 6E[\widehat{W|W>0}]$. We compute conditional ASEs for each of the different intervals. For example, the conditional ASEs of the alternative estimators for the first interval, $(E[\widehat{W|W>0}], 2E[\widehat{W|W>0}])$, are computed as follows: in (6.3), we include only pairs (a_i, p_i) where $a_i \in (E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$ and we do this for all delay estimators returning a conditional ASE for each, in that given interval. We only consider in this section models where the number of servers is large, which we choose to be $s \geq 100$. We do this here because we are more interested in the case of large number of servers, which better captures the real-life settings we are interested in, such as medium to large sized call centers. We vary the traffic intensities considered and compute interval and point estimates of the conditional ASEs in each case. In certain cases, particularly for large s, we found that our sample of actual delays collected was not large enough to infer significant statistical estimates. In that case, we simply did not display the results we obtained (e.g., see the $D/M/900$ results below).

In section 3, we classified the performances of our delay estimators as follows:

$$QL > LES \approx HOL > RCS \approx RCS - \sqrt{s} > LCS > NI$$

where " $>$ " is to be read as "performs better than" and " $\approx$ " is to be read as "performs nearly the same but slightly better than". This order of performances held for all models considered, in the range of traffic intensities considered. In this section, we see interesting results that were not captured in the overall comparisons of section 3. We report our results for the $M/M/s$ model in Tables 31-41, those for the $D/M/s$ model in Tables 42-50 and those for the $H_2/M/s$ model in Tables 51-61. Each set of tables just mentioned is followed by the corresponding figures plotting the corresponding conditional ASEs. The first observation we make is about the performance of the No-Information Steady-State estimator (NI) when studying conditional ASEs as opposed to the unconditional ASEs of the previous section. NI performs sometimes better than all other delay estimators (except the Full-Information Queue Length Estimator – QL), when the delays considered fall in the $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$ interval and the traffic intensities considered are low (e.g., $\rho = 0.9$ in the $M/M/100$ queue). This result is interesting in that it shows that it is not always necessary to use more state information in order to give better delay estimates for current waiting customers. As we consider intervals for longer delays, however, the performance of NI deteriorates (see, for example, Figure 1 for the $M/M/100$ model) .

In general, when we consider ρ large enough, and restrict our attention to long actual customer delays observed, the classification of our delay estimators is similar to that of section 3. We once more see that LES and HOL's performances are very similar. We also see that RCS and $RCS - \sqrt{s}$ are very close, with RCS performing slightly better. LCS and NI fall behind constantly, except in the case mentioned above. Finally, QL performs always best among all of our delay estimators.

In a nutshell, to draw conclusions on sections 3 and 4, we can say that given enough information about the system, the QL delay estimator is the most efficient among the estimators considered. Closely behind fall LES and HOL. In third place, we put RCS and $RCS - \sqrt{s}$. Finally, LCS and NI perform the worse with NI falling behind LCS. But, our aim in this supplement is not to determine which delay estimator to use in practice. In fact, the answer to this question lies beyond the scope of this work as it involves other parameters such as the costs involved and the preferences of the system managers. Also, as we just saw, the performances of our delay estimators depend on the length of the delays observed and a clear-cut answer is not always possible. What we do here is merely quantify the performances of alternative delay estimators that we find to be possible candidates for delay estimations in real-life systems. We study these performances by averaging over all delays observed in section 3 and then look more closely at these performances given the level of delays observed in this section.

5. Simulation Experiments for the Efficiency of the LES Delay Estimator

{secLESEff}

In this section we present simulation experiments that test the accuracy of approximations for the mean squared error of the LES delay estimator. Note that we will use HOL and LES interchangeably since we use one to approximate the other. Simulation has shown that their performances are indeed nearly identical in all cases. When the interarrival times in our model have a non-lattice cumulative distribution function, we use theorem 4.2 of the main paper and test the following approximation:

$$E[MSE(LES)] \approx (1-\rho)^2 w^2 + \frac{(2\rho-1)c_a^2 + 4\rho - 3)w}{s} + \frac{K}{s^2} \quad (5.1) \quad \{MSE\_LES\}$$

where K is given by:

$$K = \frac{3c_a^4}{2} + 4c_a^2 + \frac{9}{2} - \frac{2v_a^3}{3} \quad (5.2) \quad \{Eqn\_K\}$$

In (5.1), w denotes the delay estimate given under the LES delay estimation scheme. We replace w in what follows by our simulation point estimate for $E[W|W>0]$ and w^2 by our simulation point estimate for $E[W^2|W>0]$ where $[W|W>0]$ denotes a random variable with the distribution of the steady state waiting time conditional on the wait being positive. The reader interested in the derivation of this approximation is referred to section 4 of the main paper. This approximation is useful because it allows us to roughly quantify the performance of the LES (and hence, HOL) delay estimator, in the $GI/M/s$ queueing model, when the general renewal arrival process has non-lattice interrenewal-time distribution. The approximation should work best when $sw \to \infty$.

For deterministic interarrival times (which have a lattice distribution), we propose testing the following approximation:

$$E[MSE(LES)] \approx (1-\rho)^2 w^2 + \frac{\rho w + 2/s}{s} \quad (5.3) \quad \{MSE\_LES\_2\}$$

We now explain the derivation of this approximation. Suppose that we have deterministic interarrival times. Let $A = \{A(t) : t \geq 0\}$ be the renewal counting process associated with the deterministic interarrival times. Then, ignoring the error term, we can write: $A(w) \approx \rho s w$. Now, from equation (4.3) in the main paper, we get that:

$$var(W_{HOL}(w)) = (\rho s w + 2) \times \frac{1}{s^2} \quad (5.4) \quad \{oui1\}$$

Further, noting that $E[W_{HOL}(w)] = \frac{E[A(w)+2]}{s} = \frac{\rho s w+2}{s}$:

$$MSE(LES) = E[(W_{HOL}(w) - w)^2] = E[((W_{HOL}(w) - \frac{\rho s w + 2}{s}) + (\frac{\rho s w + 2}{s} - w))^2] \quad (5.5) \quad \{oui2\}$$

yielding approximately:

$$E[MSE(LES)] \approx (1-\rho)^2 w^2 + \frac{\rho w + 2/s}{s} \quad (5.6) \quad \{MSE\_LES\_2\}$$

The LES delay estimator is the delay (before starting service) of the last customer to have entered service, prior to our customer's arrival. This estimator is appealing because it is relatively easy to obtain and interpret. Its implementation does not require knowledge of the queueing model's parameters (number of servers, arrival process and service distributions ... etc).

We propose testing the above approximations while varying two parameters: the number of servers, s, and the traffic intensity in the system, ρ. We are particularly interested in the case of large s since we are motivated by applying this research to call centers, but we will include smaller values of s as well. We expect the approximation to be more accurate when the number of servers is large and/or the delays experienced by waiting customers are long. In particular, we consider the values: $s = 1, 10, 100, 400$ and 900. We test approximation (5.6) with exponential and H_2 ($c_a^2 = 4$) interarrival time distributions. We test approximation (5.1) with deterministic interarrival time distribution. The simulation results reported throughout are usually based on 10 independent replications of about 5 million events each, where each event is either an arrival or a service completion. We do make the runs longer when the traffic intensity is higher, or when the variability in the arrival process is high such as with H_2 arrivals (7 - 10 million events). Our selection criterion for the simulation run length is based solely on our observations when performing the simulations. When we detect high variability in the simulation estimates across runs for given s and ρ, we increase the simulation run length so as to reduce this variability.

We report our results for the $M/M/s$ model in Tables 62 - 22, those of the $D/M/s$ model in Tables 67 - 71 and those of the $H_2/M/s$ model in Tables 72 - 76. In all tables, we consider $s = 1, 10, 100, 400$ and 900. We also vary the traffic intensity $\rho = \frac{E[V]}{sE[U]}$ and report results for each value considered. In each case, we construct 95 percent confidence intervals for the estimators and report these intervals. To capture the accuracy of the approximation, we compute the relative percent difference (RPD) between ASE(LES) and the corresponding numerical approximation. We define the RPD as:

$$RPD \equiv \frac{ASE(LES) - approx}{approx} \times 100 \qquad (5.7) \quad \{\texttt{RelDiff}\}$$

where *approx* denotes the corresponding numerical approximation. We report values for all RPDs in the tables below. Studying the reported RPD values allows us to assess the accuracy of the proposed approximation. Studying the values of the reported RPDs, we immediately see that approximation (5.1) works well in the $M/M/s$ and $H_2/M/s$ queues. In these two models, we see that the absolute values of the RPDs reported are always less than 5 percent (except for the $M/M/900$ case with $\rho = 0.99$ where the $RPD \approx 7$ percent). On the other hand, approximation (5.6) performs well in the $D/M/s$ queue, except when the loads are light and s is small (e.g., $s = 10, \rho = 0.85$). Otherwise, the RPDs reported in the $D/M/s$ model are consistently close to 5 percent, sometimes reaching a remarkably low value (e.g., see the $D/M/400$ model when $\rho = 0.99$. In general, both approximations work best when the number of servers is large and the value of ρ is high.

6. Relative Efficiency of HOL Compared to QL

In this section we reproduce our simulation results for the HOL and QL delay estimators, but with a different aim in mind. Here, we compare the efficiency of the QL and HOL delay estimators and test the validity of the approximation (which can be found on p.14 of the main paper):

$$\frac{c^2_{W_{HOL,s}(w)}}{c^2_{W_{Q,s}(n)}} \approx \frac{c_a^2 + 1}{\rho} \qquad (6.1) \quad \{\texttt{Approx}\}$$

which holds when sw and n are large. This approximation compares the efficiency of the QL and *refined* HOL delay estimators. The refined HOL delay estimator (which we denote here by $HOL_{refined}$) is an announcement of ρw where w is the elapsed delay of the customer who is at the head of the line when our customer arrives to the system and joins the line of waiting customers.

We stop first on approximation (6.1). The analysis leading to this approximation can be found on pages $12-14$ of the main paper. s denotes the number of servers in our model. c_a^2 is the squared coefficient of variation (SCV) of the arrival process. ρ is the traffic intensity in the system. $W_{HOL,s}(w)$ is a random variable with the conditional distribution of the waiting time of a new arrival given that the new arrival must wait some positive amount of time, that there already is at least one customer in queue and that the customer at the head of the line has already spent time w in queue. $W_{Q,s}(n)$ is a random variable with the conditional distribution of the delay of a new arriving customer, given that the arriving customer must wait before starting service and that the queue length at the time of arrival is n. Consider the QL and refined HOL estimators. Then, the mean square errors (MSEs) for these estimators coincide with the variances $var(W_{Q,s}(n))$ and $var(W_{HOL,s}(w))$, respectively. Further, the relative mean square errors (RMSE $\equiv MSE/Mean^2$) for these estimators coincide with the respective SCVs. In this section, we test the validity of approximation (6.1) by noting that:

$$\frac{MSE(HOL_{refined})}{MSE(QL)} \approx \frac{RMSE(HOL_{refined})}{RMSE(QL)} = \frac{c^2_{W_{HOL,s}(w)}}{c^2_{W_{Q,s}(n)}} \approx \frac{c_a^2 + 1}{\rho} \qquad (6.2) \quad \{\texttt{leapOfFaith}\}$$

To approximate the MSEs of the two candidate delay estimators, we use simulation to compute the corresponding average square errors (ASEs). For large samples, the ASE should agree with the MSE in steady state. Recall that the ASE of an estimator is given by:

$$ASE \equiv \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2, \qquad (6.3) \quad \{\texttt{ASE}\}$$

where a_i is the actual delay observed and p_i is the predicted (estimated) delay. Under heavy loads, $\rho w \uparrow w$ as $\rho \uparrow 1$. There should thus be little difference between the direct and refined HOL delay estimators, in the heavily loaded systems that we consider. We therefore consider the direct HOL delay estimator in this section (announcement of w instead of ρw) and expect that approximation (6.1) above be valid for this direct HOL estimator as well.

We now present our simulation experiments. In each case, we construct 95 percent confidence intervals for the ASEs of our estimators. For a description of our simulation experiments in this supplement, please refer to section 2 of this supplement. We report our results for the $M/M/s$ model in Tables 77 - 81. We report our results for the $D/M/s$ model in Tables 82 - 86. We report our results for the $H_2/M/s$ model in Tables 87 - 91. We consider $s = 1, 10, 100, 400$ and 900. We also vary the traffic intensity $\rho = \frac{E[V]}{sE[U]}$ and report results for each value considered. The half widths of the confidence intervals reported range from less than 1 percent (high loads, small number of servers) to close to 10 percent (lighter loads, large number of servers as when $s = 900$ and $\rho = 0.93$). This is so because under lighter loads, especially when the number of servers is large, there is less data because fewer customers have to wait before beginning service. Also included in these tables are the approximation values $\frac{c_a^2+1}{\rho}$. To capture the accuracy of the approximation, we compute the relative percent difference (RPD) between ASE(HOL)/ASE(QL) and the corresponding numerical approximation. We define the RPD as:

$$RPD \equiv \frac{ASE(HOL)/ASE(QL) - approx}{approx} \times 100 \qquad (6.4) \quad \{\texttt{RelDiff}\}$$

where *approx* denotes the corresponding numerical approximation. We report values for all RPDs in the tables below. Studying the reported RPD values allows us to assess the accuracy of the proposed approximation. We see that approximation (6.1) is most accurate in the $M/M/s$ model where the RPD is consistently lower than 2 percent. It performs worse in the $D/M/s$ model where the reported RPDs range from about 2 percent to about 20 percent. Finally, the approximation performs worst in the $H_2/M/s$ model where the reported RPDs range from 5 percent to over 30 percent. In all cases however, we note that our approximation is indeed a useful one, particularly when the traffic intensity considered in the system is high.

7. The Effect of Delay Information: RCS$-f(s)$

{secRCSEffect

In this section, we consider a family of delay estimators, RCS-$f(s)$, for functions $f(s) = s, 4\sqrt{s}, 2\sqrt{s}, \sqrt{s}$ and $log(s)$. RCS-$f(s)$ announces the delay of the most recent arrival among the last $f(s)$ customers to have completed service. We use simulation to study the differences between these delay estimators. We wish here to quantify the impact of using past delay information: we expect that increasing the amount of information used would lead to an improvement in the performance of the estimator. This improvement of performance is quantified via a decrease in the value of the average squared error reported. However, using more information implies more data processing which could prove to be costly. We report here simulation results showing that using the delay information of all customers who have completed service is not necessary. Indeed, from heavy-traffic analysis, we deduce that the most recent arrival time of a customer that has completed service is very likely to occur among the last $c\sqrt{s}$ customers when s is large. Here, we use simulation to find that value of c, which turns out to be equal to 4.

Under some circumstances, the LCS (Last Customer to Complete Service) and LES (Last Customer to Enter Service) delay estimators will be similar, but they actually can be very different when the number of servers s is large. To see why, assume that s is large; let X denote an exponential random variable with mean $E[X] = 1/s$ and V be a generic service time random variable, i.e., an exponential random variable with mean $E[V] = 1$. Further, assume that all servers are busy upon arrival of our customer, at time t, and that she found at least one other customer waiting in line ahead of her when she arrived. These assumptions are not unreasonable in the heavily loaded systems that we consider. The last customer to have completed service prior to t must have experienced her delay before time $t - V$. The last customer to have entered service prior to t must have experienced her delay at most X time units ago. This is so because the random times between successive service completions in the $GI/M/s$ model are exponential random variables with mean $1/s$. When s is large, this means that the last customer to complete service may have experienced her delay much before the last customer to enter service. This explains why LCS and LES can be different, especially when s is large, since more recent information is more relevant to the current state of the system.

We are thus lead to propose other candidate delay estimators based on the delay experience of customers that have already completed service. The first is the delay experienced by the customer that arrived most recently (and thus entered service most recently, under our First

In First Out service discipline), among those customers who have already completed service (RCS). A disadvantage of the RCS estimator is that we must analyze a lot of data, going arbitrarily far back in the past. From heavy-traffic analysis, we deduce that the most recent arrival time of a customer that has completed service is very likely to occur among the last $c\sqrt{s}$ customers when s is large. Simulation shows that the corresponding value of c is 4. We find the same value of c for all values of ρ, s and interarrival time distributions considered. The reader interested in this heavy-traffic analysis is referred to Section 7 of the main paper (pages 24-25). So, we introduce another estimator which requires less information processing: RCS-$c\sqrt{s}$ is the delay of the customer to have arrived most recently among the last $c\sqrt{s}$ customers who have already completed service.

In the tables below, we report the *average square errors* (ASE) for each considered RCS-$f(s)$ delay estimator, along with corresponding 95 percent confidence intervals. These confidence intervals and point estimates are based on 10 independent simulation runs of length 1 million events each. Our simulation run length in this section is shorter than in other sections of the supplement because our aim here is determine the value of c; this value of c is not expected to change when we increase the simulation run length. We only consider values of $s \geq 100$ since we don't expect to see substantial differences between our estimators when the number of servers is small. In our tables, we include in parentheses next to the reported ASE point estimate, the value of relative percent difference (RPD). We define the RPD as:

$$RPD \equiv \frac{ASE(RCS - f(s)) - ASE(RCS)}{ASE(RCS)} \times 100 \qquad (7.1)$$ {RelDiff}

Studying the values of the reported RPDs allows us to assess the impact of the amount of delay information used. The smaller the RPD, the closer the performance of our estimator is to the best possible estimator here, RCS. The following observations are true throughout this section. We see that the performance of RCS-$2\sqrt{s}$ is very close to that of RCS; the RPDs reported are less than 1 percent for all models considered. RCS-$\sqrt{s}$ falls slightly behind but is still very similar to RCS. The RPDs reported in this case are consistently less than 10 percent. Finally, RCS-$log(s)$ performs the worse, nearly as bad as LCS. That is consistent with our analysis, since $log(s) \leq 3$ for all values of s considered. The amount of delay information processed in this case is actually too little.

We report our results for the $M/M/s$ model in Tables 92 - 94, those for the $H_2/M/s$ model in Tables 95 - 97 and those for the $D/M/s$ model in Tables 98 - 100.

8. Distributions of The Delay Estimators

{secDistHOL}

In this section, we look more carefully at the distribution of actual delays observed given that delay estimates announced are in a small interval about a certain value which we choose. Our first aim is to test approximation (4.13) of the main paper. This approximation states that:

$$W_{HOL}(w) \approx N(\rho w, \rho w(c_a^2 + 1)/s) \qquad (8.1)$$ {normapprox}

Recall that $W_{HOL}(w)$ denotes a random variable with the conditional distribution of the waiting time (before starting service) of a new arrival given that the new arrival must join the queue, given that there is already at least one customer in queue, and given that the customer at the head of the line has already spent time w in queue. That is, w is the delay estimate given according to the HOL delay estimation scheme. We expect this approximation to hold as $sw \to \infty$. Thus, in order to test it, we only consider models with a large number of servers (here, we consider $s = 100$), large w (here, we consider $w \approx 2E[W|W > 0]$) and high ρ (here, $\rho = 0.95$). Recall that $[W|W > 0]$ denotes a random variable with the conditional distribution of the steady state waiting time given that the wait is positive.

We proceed as follows: we simulate the model under consideration and collect the actual delays observed when the delay estimate given according to the HOL scheme, w, is approximately equal to $2E[W|W > 0]$. By approximately equal we mean that w should be in a small interval about $2E[W|W > 0]$ (or $2E[\widehat{W|W} > 0]$, our simulation point estimate for this quantity). We choose this small interval so as to have a large enough sample size for the data thus collected. We then plot a histogram for the actual delays collected. This histogram describes the distribution of $W_{HOL}(w)$ for the chosen w in the given model, thus enabling us to test (8.1) above. In Figures 62, 67 and 72, we display histogram plots for the $M/M/100$, $D/M/100$ and $H_2/M/100$ models, respectively. To better assess how well the Normal distribution fits the data distribution, we plot a Normal curve with the same mean and variance on top of our histograms. We see in these figures that the distribution of $W_{HOL}(w)$ is indeed very close to being a Normal distribution. In the $M/M/100$ model, with $\rho = 0.95$, we can compute $E[W|W > 0]$ analytically and get and exact value of: $E[W|W > 0] = 0.2$. Our simulation point estimate for $E[W_{HOL}(w)]$ based on our sample of actual delays observed is $E[\widehat{W_{HOL}}(w)] \approx 0.4003$, which is slightly larger than $\rho w \approx \rho \times 2E[W|W > 0] \approx 0.38$. Our simulation point estimate for the variance is $var[\widehat{W_{HOL}}(w)] \approx 0.008$ while the variance $var[W_{HOL}(w)]$ in (8.1) yields 0.0076 in this case. Note that approximations (4.7) and (4.9) of the main paper for the mean and variance of $W_{HOL}(w)$ yield respectively 0.4 and 0.0078 which are closer to our simulation values. For completeness, we restate these approximations:

$$sE[W_{HOL,s}(w)] - \rho sw \rightarrow \frac{(c_a^2 + 3)}{2} \ , \qquad (8.2) \quad \{a4\}$$

and,

$$s^2 Var(W_{HOL,s}(w)) - \rho sw(c_a^2 + 1) \rightarrow \left(\frac{5(c_a^2 + 1)^2}{4} - \frac{2\nu_a^3}{3} + 1 \right) \qquad (8.3) \quad \{a4c\}$$

Hence, we see that the simulation estimates reported agree quite closely with the above approximation. In the $D/M/100$ model with $\rho = 0.95$, our simulation point estimate for $E[W|W > 0]$ is given by: $E[\widehat{W|W > 0}] \approx 0.1$. From our simulations we see that: $E[\widehat{W_{HOL}(w)}] \approx 0.1996$ whereas $\rho w \approx \rho \times 2E[\widehat{W|W > 0}] \approx 0.19$. On the other hand, $var[\widehat{W_{HOL}(w)}] \approx 0.002$ from simulations, and is given by 0.0019 in (8.1). Once more, we see that the simulation estimates reported agree quite closely with the above approximation. Note that approximations (4.7) and (4.9) of the main paper for the mean and variance of $W_{HOL}(w)$ yield respectively 0.205 and 0.0021 which are closer to our simulation values.

Finally, in the $H_2/M/100$ model, we see from simulations that $E[\widehat{W|W > 0}] \approx 0.48$ when $\rho = 0.95$. This yields from (8.1) that $E[W_{HOL}(w)]$ should be approximately equal to $\rho w \approx 0.912$. From our simulations, we see that $E[\widehat{W_{HOL}(w)}] \approx 0.9625$. On the other hand, (8.1) yields $var[W_{HOL}(w)] \approx 0.0456$ and our simulation point estimate is $var[\widehat{W_{HOL}(w)}] \approx 0.0448$. Note that approximations (4.7) and (4.9) of the main paper for the mean and variance of $W_{HOL}(w)$ yield respectively 0.947 and 0.0448 which are closer to our simulation values (remarkably so for the variance).

Thus, we see that our approximation is indeed a valid one in the three models considered above. Based on our simulation results, we see that using the following alternative approximation yields sightly better results:

$$W_{HOL}(w) \approx N(\rho w + \frac{c_a^2 + 3}{2s}, \rho w \frac{c_a^2 + 1}{s} + \frac{5(c_a^2 + 1)^2}{4s^2} - \frac{2\nu_a^3}{3s^2} + \frac{1}{s^2}) \qquad (8.4) \quad \{normapprox2\}$$

In this section, we also go beyond studying the distribution of $W_{HOL}(w)$ and include histograms for $W_{LES}(w)$ and $W_{RCS}(w)$. Recall that $W_{LES}(w)$ denotes a random variable with the conditional distribution of the waiting time (before starting service) of a new arrival given that the new arrival must join the queue, given that the last customer to have entered service prior to the new arrival had delay w. Similarly, $W_{RCS}(w)$ denotes a random variable with the conditional distribution of the waiting time (before starting service) of a new arrival given that the new arrival must join the queue, given that the most recent arrival among all customers that have completed service had delay w. Once more, we restrict our w to being approximately equal to $2E[W|W > 0]$. We show the corresponding histograms in Figures 65-66 for the $M/M/100$ model, 69-70 for the $D/M/100$ model and 75-76 for the $H_2/M/100$ model. To plot the histogram in the LES case for example, we collect the actual delays observed when the delay estimate given according to the LES delay estimation scheme is approximately equal to $2E[W|W > 0]$ (or our simulation point estimate for this quantity, $2E[\widehat{W|W >} 0]$). We then plot a histogram for the actual delays thus collected, thus plotting the distribution of $W_{LES}(w)$. We do the same for $W_{RCS}(w)$. We see that our simulation point estimates for the mean and variance of $W_{LES}(w)$ and $W_{RCS}(w)$ are quite close to those of $W_{HOL}(w)$ (see for example Figures 65 and 66), and we once more see an approximately Normal distribution.

Finally, we also plot histograms for the distributions of the LES and RCS delay estimations given when the delay estimate given according to the HOL delay estimation scheme, w, is approximately equal to $2E[W|W > 0]$. We see throughout that the variance of the LES and RCS estimations is considerably smaller than that of $W_{HOL}(w)$ (e.g., see Figure 63). In the $M/M/100$ and $H_2/M/100$ models (e.g., see Figures 64 and 74), we see that the LES and RCS predictions given are normally distributed. In the $D/M/100$ model (see Figure 68), the distribution of the LES estimations thus collected is not Normal. At present, we did not investigate the shape of the distribution observed in this case.

9. Conclusions

In this supplement to the main paper we presented simulation results investigating the performances of 6 delay estimators: (i) the delay of the last customer to enter service (LES), (ii) the delay experienced so far by the customer at the head of the line (HOL), (iii) the delay experienced by the last customer to complete service (LCS), (iv) the delay experienced by the customer to have arrived most recently among those who have completed service (RCS), (v) the delay of our customer in the $GI/M/s$ model assuming full information at the arrival epoch (QL) and (vi) the steady state mean waiting time, assuming no information beyond the model (NI). We concluded that LES and HOL are very similar, with both being more accurate than others. For large s, RCS is far superior to LCS, because customers need not complete service in the same order they arrive. QL outperforms all other delay estimators but has the disadvantage of assuming full knowledge about the model at the arrival epoch. NI almost always falls behind, except when we restrict our attention to low traffic intensities and relatively short delays. We saw that simulation provided additional support to some theoretical results of the main paper thus further validating them. These results relate mostly to the performance of the LES delay estimator (or equivalently, HOL) in the $GI/M/s$ model. We also used simulation to gain insight into the performances of our delay estimators when we didn't have corresponding theoretical results. We did this, for instance, when studying the performances of the RCS and LCS delay estimators. We went beyond studying the overall performances of our estimators and studied differences in performances when restricting attention to specific delay levels in the system. We then saw that these performances can indeed depend on the level of delays considered.

It is important to note finally that the simulation results of this supplement merely aim at quantifying the differences between the performances of the alternative delay estimators that we proposed. In practice, there are a number of issues that we do not consider here, such as the amount of information available to the system manager and the costs involved in implementing the different delay estimation schemes. The question of which delay estimator to use in a real-life problem cannot be answered without incorporating all of these factors.

10. Tables and Figures

ASE *in the* $M/M/s$ *model with* $s=1$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	20.09	42.24	42.35	44.09	44.09	44.09	405.0
	±0.417	±0.766	±0.785	±0.776	±0.776	±0.776	±23.4
0.93	14.36	30.56	30.72	32.36	32.36	32.36	207.5
	±0.186	±0.371	±0.385	±0.373	±0.373	±0.373	±10.4
0.9	9.989	21.76	21.96	23.46	23.46	23.46	100.6
	±0.084	±0.186	±0.206	±0.192	±0.192	±0.192	±3.44
0.85	6.678	15.07	15.38	16.63	16.63	16.63	44.94
	±0.043	±0.093	±0.0951	±0.010	±0.010	±0.010	±0.883

Table 1: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s=1$ and $\mu=1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $M/M/s$ *model with* $s=1$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.95	0.05003	0.1052	0.1054	0.1098	0.1098	0.1098	1.008
0.93	0.06990	0.1488	0.1495	0.1575	0.1575	0.1575	1.010
0.9	0.1002	0.2183	0.2203	0.2354	0.2353	0.2354	1.009
0.85	0.1497	0.3379	0.3448	0.3727	0.3727	0.3727	1.008

Table 2: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s=1$ and $\mu=1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE *in the* $M/M/s$ *model with* $s = 10$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	0.4954	1.005	1.005	1.075	1.088	1.193	25.72
	±0.0232	±0.0416	±0.0413	±0.0412	±0.0423	±0.0412	±4.81
0.95	0.1980	0.4162	0.4171	0.4831	0.4940	0.5869	3.9609
	±0.00251	±0.00399	±0.00416	±0.00394	±0.00412	±0.00413	±0.228
0.93	0.1421	0.3033	0.3046	0.3672	0.3774	0.4621	2.0014
	±0.00133	±0.00322	±0.00374	±0.00357	±0.00330	±0.00360	±0.0659
0.9	0.1003	0.2185	0.2204	0.2792	0.2884	0.3629	1.0099
	±0.00168	±0.00329	±0.00416	±0.00359	±0.00346	±0.00362	±0.0490
0.85	0.0661	0.1499	0.1528	0.2037	0.2112	0.2692	0.4409
	±0.000324	±0.000757	±0.00124	±0.000924	±0.000854	±0.000970	±0.00825

Table 3: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $M/M/s$ *model with* $s = 10$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS - \sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.0205	0.0417	0.0417	0.0446	0.0451	0.0495	1.067
0.95	0.05021	0.1055	0.1057	0.1225	0.1252	0.1488	1.004
0.93	0.07042	0.1503	0.1510	0.1820	0.1870	0.2290	0.9918
0.90	0.09994	0.2176	0.2195	0.2781	0.2873	0.3615	1.006
0.85	0.1495	0.3387	0.3455	0.4604	0.4775	0.6085	0.9966

Table 4: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

			ASE in the M/M/s model with $s = 100$				
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	5.033×10^{-3}	1.023×10^{-2}	1.023×10^{-2}	1.252×10^{-2}	1.287×10^{-2}	2.660×10^{-2}	2.549×10^{-1}
	$\pm 2.46 \times 10^{-4}$	$\pm 5.04 \times 10^{-4}$	$\pm 5.09 \times 10^{-4}$	$\pm 5.11 \times 10^{-4}$	$\pm 5.14 \times 10^{-4}$	$\pm 6.17 \times 10^{-4}$	$\pm 3.55 \times 10^{-2}$
0.95	2.041×10^{-3}	4.258×10^{-3}	4.269×10^{-3}	6.380×10^{-3}	6.674×10^{-3}	1.652×10^{-2}	4.180×10^{-2}
	$\pm 4.22 \times 10^{-5}$	$\pm 5.85 \times 10^{-5}$	$\pm 6.028 \times 10^{-5}$	$\pm 6.53 \times 10^{-5}$	6.96×10^{-5}	$\pm 1.72 \times 10^{-5}$	2.71×10^{-3}
0.93	1.442×10^{-3}	3.069×10^{-3}	3.084×10^{-3}	5.058×10^{-3}	5.317×10^{-3}	1.306×10^{-2}	2.083×10^{-2}
	$\pm 1.66 \times 10^{-5}$	$\pm 2.53 \times 10^{-5}$	$\pm 2.63 \times 10^{-5}$	$\pm 2.71 \times 10^{-5}$	$\pm 3.29 \times 10^{-5}$	$\pm 1.27 \times 10^{-4}$	$\pm 1.22 \times 10^{-3}$
0.90	9.940×10^{-4}	2.168×10^{-3}	2.185×10^{-3}	3.949×10^{-3}	4.157×10^{-3}	9.379×10^{-3}	9.702×10^{-3}
	$\pm 2.97 \times 10^{-5}$	$\pm 5.87 \times 10^{-5}$	$\pm 5.96 \times 10^{-5}$	$\pm 8.38 \times 10^{-5}$	$\pm 9.11 \times 10^{-5}$	$\pm 2.74 \times 10^{-4}$	$\pm 7.06 \times 10^{-4}$

Table 5: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $M/M/s$ *model with* $s = 100$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS - \sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.01982	0.04031	0.04033	0.04931	0.05075	0.1053	1.004
0.95	0.04961	0.1042	0.1043	0.1551	0.1624	0.4022	1.016
0.93	0.07001	0.1493	0.1506	0.2465	0.2584	0.6342	1.011
0.9	0.1013	0.2203	0.2226	0.4002	0.4226	0.9512	0.9846

Table 6: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE *in the* $M/M/s$ *model with* $s = 400$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	3.145×10^{-4}	6.424×10^{-4}	6.427×10^{-4}	9.273×10^{-4}	9.706×10^{-4}	4.032×10^{-3}	1.548×10^{-2}
	1.35×10^{-5}	2.46×10^{-5}	2.50×10^{-5}	2.36×10^{-5}	2.35×10^{-5}	5.36×10^{-5}	2.05×10^{-3}
0.95	1.262×10^{-4}	2.646×10^{-4}	2.651×10^{-4}	5.168×10^{-4}	5.496×10^{-4}	2.168×10^{-3}	2.565×10^{-3}
	5.51×10^{-6}	1.11×10^{-5}	1.13×10^{-5}	1.43×10^{-5}	1.49×10^{-5}	8.94×10^{-5}	3.25×10^{-4}
0.93	9.109×10^{-5}	1.951×10^{-4}	1.960×10^{-4}	4.263×10^{-4}	4.523×10^{-4}	1.542×10^{-3}	1.39×10^{-3}
	3.77×10^{-6}	7.18×10^{-6}	7.35×10^{-6}	9.72×10^{-6}	1.02×10^{-5}	$7.80 \times 10-5$	1.83×10^{-4}
0.9	6.451×10^{-5}	1.401×10^{-4}	1.414×10^{-4}	3.486×10^{-4}	3.704×10^{-4}	9.574×10^{-4}	6.358×10^{-4}
	4.43×10^{-6}	7.31×10^{-6}	7.77×10^{-6}	1.21×10^{-5}	1.39×10^{-5}	9.41×10^{-5}	1.32×10^{-4}

Table 7: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

	$RASE$ *in the* $M/M/s$ *model with* $s = 400$						
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.01959	0.04002	0.04004	0.05777	0.06046	0.2512	0.9643
0.95	0.04945	0.1037	0.1039	0.2025	0.2154	0.8495	1.005
0.93	0.06724	0.1440	0.1447	0.3135	0.3339	1.1380	1.0290
0.9	0.1007	0.2187	0.2207	0.5440	0.5781	1.494	0.9922

Table 8: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

	ASE in the $M/M/s$ model with $s = 900$						
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.320×10^{-4}	2.668×10^{-4}	2.669×10^{-4}	3.542×10^{-4}	3.680×10^{-4}	1.923×10^{-3}	1.513×10^{-2}
	8.86×10^{-6}	1.85×10^{-5}	1.86×10^{-5}	1.85×10^{-5}	1.87×10^{-5}	2.93×10^{-5}	3.44×10^{-3}
0.98	6.334×10^{-5}	1.286×10^{-4}	1.287×10^{-4}	2.123×10^{-4}	2.252×10^{-4}	1.381×10^{-3}	3.244×10^{-3}
	2.15×10^{-6}	4.24×10^{-6}	4.25×10^{-6}	4.43×10^{-6}	4.53×10^{-6}	2.82×10^{-5}	3.26×10^{-4}
0.95	2.484×10^{-5}	5.219×10^{-5}	5.229×10^{-5}	1.246×10^{-4}	1.340×10^{-4}	6.114×10^{-4}	4.934×10^{-4}
	1.79×10^{-6}	3.85×10^{-6}	3.95×10^{-6}	5.11×10^{-6}	5.49×10^{-6}	6.25×10^{-5}	1.018×10^{-4}
0.93	1.752×10^{-5}	3.821×10^{-5}	3.834×10^{-5}	1.042×10^{-4}	1.101×10^{-4}	3.701×10^{-4}	2.342×10^{-4}
	1.94×10^{-6}	4.09×10^{-6}	4.11×10^{-6}	8.59×10^{-6}	9.41×10^{-6}	6.83×10^{-5}	6.43×10^{-5}

Table 9: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $M/M/s$ *model with* $s = 900$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.99	0.009346	0.01871	0.01883	0.02494	0.02594	0.1351	1.063
0.98	0.01962	0.03971	0.03984	0.06536	0.06952	0.4253	0.9997
0.95	0.04973	0.1045	0.1047	0.2495	0.2684	1.225	0.9882
0.93	0.07084	0.1541	0.1547	0.4187	0.4447	1.495	0.9451

Table 10: A comparison of the efficiency of different real-time delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

			ASE in the $D/M/s$ *model with* $s = 1$				
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	10.13 ±0.148	11.61 ±0.146	11.61 ±0.157	12.56 ±0.148	12.56 ±0.148	12.56 ±0.148	101.1 ±7.23
0.93	7.322 ±0.0805	8.791 ±0.0780	8.794 ±0.0862	9.731 ±0.0804	9.731 ±0.0804	9.731 ±0.0804	52.73 ±2.42
0.9	5.192 ±0.0377	6.644 ±0.0372	6.647 ±0.0409	7.556 ±0.0397	7.556 ±0.0397	7.556 ±0.0397	26.76 ±0.942
0.85	3.528 ±0.0183	4.955 ±0.0178	4.954 ±0.0203	5.819 ±0.0203	5.819 ±0.0206	5.819 ±0.0203	12.43 ±0.357

Table 11: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $D/M/s$ *model with* $s = 1$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.95	0.09916	0.1136	0.1136	0.1230	0.1230	0.1230	0.9899
0.93	0.1375	0.1650	0.1651	0.1827	0.1827	0.1827	0.9898
0.90	0.1936	0.2478	0.2479	0.2818	0.2818	0.2818	0.9979
0.85	0.2845	0.3995	0.3995	0.4692	0.4692	0.4692	1.002

Table 12: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE in the D/M/s model with s = 10							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	2.485×10^{-1}	2.633×10^{-1}	2.633×10^{-1}	2.990×10^{-1}	3.050×10^{-1}	3.572×10^{-1}	5.933
	$\pm 8.36 \times 10^{-3}$	$\pm 8.34 \times 10^{-3}$	$\pm 8.61 \times 10^{-3}$	$\pm 8.51 \times 10^{-3}$	$\pm 8.57 \times 10^{-3}$	$\pm 8.55 \times 10^{-3}$	± 1.02
0.95	1.010×10^{-1}	1.158×10^{-1}	1.157×10^{-1}	1.498×10^{-1}	1.552×10^{-1}	2.004×10^{-1}	1.011
	$\pm 1.76 \times 10^{-3}$	$\pm 1.79 \times 10^{-3}$	$\pm 2.03 \times 10^{-3}$	$\pm 1.92 \times 10^{-3}$	$\pm 1.88 \times 10^{-3}$	$\pm 1.91 \times 10^{-3}$	$\pm 8.27 \times 10^{-2}$
0.93	7.301×10^{-2}	8.759×10^{-2}	8.773×10^{-2}	1.207×10^{-1}	1.257×10^{-1}	1.664×10^{-1}	5.240×10^{-1}
	$\pm 1.04 \times 10^{-3}$	$\pm 1.05 \times 10^{-3}$	$\pm 1.32 \times 10^{-3}$	$\pm 1.22 \times 10^{-3}$	$\pm 1.14 \times 10^{-3}$	$\pm 1.22 \times 10^{-3}$	$\pm 2.89 \times 10^{-2}$
0.9	5.182×10^{-2}	6.630×10^{-2}	6.632×10^{-2}	9.769×10^{-2}	1.021×10^{-1}	1.365×10^{-1}	2.658×10^{-1}
	$\pm 5.77 \times 10^{-4}$	$\pm 5.70 \times 10^{-4}$	$\pm 9.12 \times 10^{-4}$	$\pm 7.68 \times 10^{-4}$	$\pm 6.55 \times 10^{-4}$	$\pm 7.81 \times 10^{-4}$	$\pm 1.19 \times 10^{-2}$
0.85	3.519×10^{-2}	4.943×10^{-2}	4.935×10^{-2}	7.789×10^{-2}	8.137×10^{-2}	1.063×10^{-1}	1.236×10^{-1}
	$\pm 2.49 \times 10^{-4}$	$\pm 2.57 \times 10^{-4}$	$\pm 5.73 \times 10^{-4}$	$\pm 4.65 \times 10^{-4}$	$\pm 2.79 \times 10^{-4}$	$\pm 4.71 \times 10^{-4}$	$\pm 5.33 \times 10^{-3}$

Table 13: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

			$RASE$ *in the* $D/M/s$ *model with* $s = 10$				
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.04025	0.04265	0.05786	0.04843	0.04265	0.04940	0.9713
0.95	0.09902	0.1135	0.1965	0.1469	0.1134	0.1522	0.9915
0.93	0.1372	0.1647	0.3126	0.2267	0.1647	0.2361	0.9853
0.9	0.1932	0.2472	0.5089	0.3643	0.2473	0.3807	0.9918
0.85	0.2837	0.3985	0.8574	0.6280	0.3978	0.6561	0.9969

Table 14: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE in the D/M/s model with $s = 100$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	2.476×10^{-3}	2.624×10^{-3}	2.624×10^{-3}	3.773×10^{-3}	3.945×10^{-3}	1.033×10^{-2}	6.151×10^{-2}
	$\pm 5.04 \times 10^{-5}$	$\pm 5.02 \times 10^{-5}$	$\pm 5.14 \times 10^{-5}$	$\pm 5.27 \times 10^{-5}$	$\pm 5.35 \times 10^{-5}$	$\pm 1.13 \times 10^{-4}$	3.82×10^{-3}
0.95	1.007×10^{-3}	1.154×10^{-3}	1.153×10^{-3}	2.203×10^{-3}	2.338×10^{-3}	6.381×10^{-3}	1.008×10^{-2}
	$\pm 1.70 \times 10^{-5}$	$\pm 1.78 \times 10^{-5}$	$\pm 1.84 \times 10^{-5}$	$\pm 2.73 \times 10^{-5}$	$\pm 2.88 \times 10^{-5}$	$\pm 1.21 \times 10^{-4}$	$\pm 3.96 \times 10^{-4}$
0.93	7.250×10^{-4}	8.710×10^{-4}	8.713×10^{-4}	1.848×10^{-3}	1.960×10^{-3}	4.897×10^{-3}	5.197×10^{-3}
	$\pm 1.58 \times 10^{-5}$	$\pm 1.58 \times 10^{-5}$	$\pm 1.70 \times 10^{-5}$	$\pm 2.69 \times 10^{-5}$	$\pm 2.93 \times 10^{-5}$	$\pm 1.32 \times 10^{-4}$	$\pm 3.18 \times 10^{-4}$
0.9	5.189×10^{-4}	6.65×10^{-4}	6.64×10^{-4}	1.54×10^{-3}	1.63×10^{-3}	3.44×10^{-3}	2.681×10^{-3}
	$\pm 1.52 \times 10^{-5}$	$\pm 1.56 \times 10^{-5}$	$\pm 1.67 \times 10^{-5}$	$\pm 3.50 \times 10^{-5}$	$\pm 3.71 \times 10^{-5}$	$\pm 1.52 \times 10^{-4}$	$\pm 2.30 \times 10^{-4}$

Table 15: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ in the $D/M/s$ model with $s = 100$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS - \sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.04043	0.04286	0.04285	0.06162	0.06442	0.1688	1.004
0.95	0.1000	0.1146	0.1146	0.2189	0.2323	0.6340	1.001
0.93	0.1382	0.1661	0.1662	0.3524	0.3739	0.9338	0.9910
0.90	0.1939	0.2484	0.2482	0.5766	0.6084	1.284	1.002

Table 16: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE *in the* $D/M/s$ *model with* $s = 400$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	3.035×10^{-4}	3.130×10^{-4}	3.129×10^{-4}	4.602×10^{-4}	4.831×10^{-4}	2.224×10^{-3}	1.416×10^{-2}
	$\pm 1.16 \times 10^{-5}$	$\pm 1.15 \times 10^{-5}$	$\pm 1.17 \times 10^{-5}$	$\pm 1.17 \times 10^{-5}$	$\pm 1.18 \times 10^{-5}$	$\pm 3.68 \times 10^{-5}$	$\pm 1.77 \times 10^{-3}$
0.98	1.556×10^{-4}	1.649×10^{-4}	1.648×10^{-4}	3.060×10^{-4}	3.264×10^{-4}	1.620×10^{-3}	3.895×10^{-3}
	$\pm 3.37 \times 10^{-6}$	$\pm 3.34 \times 10^{-6}$	$\pm 3.42 \times 10^{-6}$	$\pm 3.51 \times 10^{-6}$	$\pm 3.40 \times 10^{-6}$	$\pm 3.43 \times 10^{-5}$	$\pm 2.64 \times 10^{-4}$
0.95	6.329×10^{-5}	7.248×10^{-5}	7.245×10^{-5}	1.945×10^{-4}	2.094×10^{-4}	7.549×10^{-4}	6.451×10^{-4}
	$\pm 2.41 \times 10^{-6}$	$\pm 2.41 \times 10^{-6}$	$\pm 2.50 \times 10^{-6}$	$\pm 4.09 \times 10^{-6}$	$\pm 4.56 \times 10^{-6}$	$\pm 3.31 \times 10^{-5}$	$\pm 7.27 \times 10^{-5}$
0.93	4.583×10^{-5}	5.497×10^{-5}	5.511×10^{-5}	1.642×10^{-4}	1.758×10^{-4}	4.975×10^{-4}	3.325×10^{-4}
	$\pm 2.41 \times 10^{-6}$	$\pm 2.55 \times 10^{-6}$	$\pm 2.64 \times 10^{-6}$	$\pm 5.54 \times 10^{-6}$	$\pm 6.74 \times 10^{-6}$	$\pm 3.97 \times 10^{-5}$	$\pm 3.84 \times 10^{-5}$
0.9	3.063×10^{-5}	3.988×10^{-5}	3.931×10^{-5}	1.277×10^{-4}	1.372×10^{-4}	2.703×10^{-4}	1.415×10^{-4}
	$\pm 4.88 \times 10^{-6}$	$\pm 5.45 \times 10^{-6}$	$\pm 5.81 \times 10^{-6}$	$\pm 1.85 \times 10^{-5}$	$\pm 1.91 \times 10^{-5}$	$\pm 6.54 \times 10^{-5}$	$\pm 4.31 \times 10^{-5}$

Table 17: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $D/M/s$ *model with* $s = 400$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.99	0.02065	0.02129	0.02129	0.03131	0.03287	0.1508	0.9633
0.98	0.04028	0.04269	0.04268	0.07923	0.08452	0.4195	1.0087
0.95	0.1015	0.1163	0.1162	0.3121	0.3359	1.2112	1.035
0.93	0.1362	0.1634	0.1638	0.4879	0.5225	1.4786	0.9884
0.90	0.1936	0.2521	0.2485	0.8075	0.8676	1.7088	0.9948

Table 18: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE in the $D/M/s$ model with $s = 900$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	6.013×10^{-5}	6.199×10^{-5}	6.198×10^{-5}	1.056×10^{-4}	1.122×10^{-4}	7.958×10^{-4}	2.789×10^{-3}
	$\pm 2.60 \times 10^{-6}$	$\pm 2.59 \times 10^{-6}$	$\pm 2.62 \times 10^{-6}$	$\pm 2.81 \times 10^{-6}$	$\pm 2.87 \times 10^{-6}$	$\pm 1.96 \times 10^{-5}$	$\pm 3.66 \times 10^{-4}$
0.98	3.088×10^{-5}	3.272×10^{-5}	3.088×10^{-5}	7.385×10^{-5}	7.968×10^{-5}	5.277×10^{-4}	7.767×10^{-4}
	$\pm 9.77 \times 10^{-7}$	$\pm 9.82 \times 10^{-7}$	$\pm 1.00 \times 10^{-6}$	$\pm 1.24 \times 10^{-6}$	$\pm 1.32 \times 10^{-6}$	2.16×10^{-5}	6.04×10^{-5}
0.95	1.35×10^{-5}	1.53×10^{-5}	1.53×10^{-5}	5.01×10^{-5}	5.40×10^{-5}	2.148×10^{-4}	1.618×10^{-4}
	$\pm 1.50 \times 10^{-6}$	$\pm 1.47 \times 10^{-6}$	$\pm 1.52 \times 10^{-6}$	$\pm 2.57 \times 10^{-6}$	$\pm 2.89 \times 10^{-6}$	$\pm 3.53 \times 10^{-5}$	$\pm 4.97 \times 10^{-5}$
0.93	9.42×10^{-6}	1.126×10^{-5}	1.125×10^{-5}	4.175×10^{-5}	4.345×10^{-5}	1.202×10^{-4}	6.099×10^{-5}
	$\pm 1.68 \times 10^{-6}$	$\pm 1.73 \times 10^{-6}$	$\pm 1.77 \times 10^{-6}$	$\pm 6.55 \times 10^{-6}$	$\pm 6.46 \times 10^{-6}$	$\pm 3.370 \times 10^{-5}$	$\pm 1.895 \times 10^{-5}$

Table 19: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $D/M/s$ *model with* $s = 900$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.99	0.02069	0.02133	0.02132	0.03631	0.03861	0.2738	0.9596
0.98	0.03989	0.04227	0.04226	0.09540	0.10294	0.68171	1.003
0.95	0.09447	0.10680	0.10721	0.35003	0.37672	1.49886	1.0289
0.93	0.1288	0.1539	0.1539	0.5709	0.5941	1.644	0.9841

Table 20: A comparison of the efficiency of different real-time delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE in the $H_2/M/s$ model with $s = 1$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	48.66	226.4	226.5	231.1	231.1	231.1	2339
	±1.13	±5.14	±5.23	±5.15	±5.15	±5.15	±425
0.93	34.33	154.4	154.4	158.9	158.9	158.9	1151
	±0.625	±2.94	±2.94	±2.95	±2.95	±2.95	±181
0.9	23.48	101.3	101.4	105.5	105.5	105.5	552.9
	±0.366	±2.32	±2.36	±2.35	±2.35	±2.35	±103
0.85	14.95	59.99	60.15	63.89	63.89	63.89	224.4
	±0.104	±0.515	±0.530	±0.514	±0.514	±0.514	±6.20

Table 21: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $H_2/M/s$ *model with* $s=1$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.95	0.02057	0.09572	0.09572	0.09768	0.09768	0.09768	0.9889
0.93	0.02933	0.13191	0.1319	0.1358	0.1358	0.1358	0.9836
0.90	0.04260	0.1837	0.1839	0.1915	0.1915	0.1915	1.0031
0.85	0.06610	0.2652	0.2659	0.2824	0.2824	0.2824	0.9921

Table 22: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s=1$ and $\mu=1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE *in the* $H_2/M/s$ *model with* $s = 10$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.284	6.259	6.259	6.435	6.512	6.730	159.4
	±0.0690	±0.404	±0.410	±0.406	±0.405	±0.595	±25.8
0.95	0.481	2.23	2.23	2.39	2.46	2.65	22.9
	±0.00811	±0.0466	±0.0482	±0.0466	±0.0470	±0.0808	±0.906
0.93	0.3424	1.542	1.543	1.698	1.752	1.940	11.53
	±0.00691	±0.0346	±0.0370	±0.0351	±0.0349	±0.0346	0.679
0.9	0.2344	1.012	1.013	1.156	1.1799	1.370	5.436
	±0.00359	±0.0181	±0.0203	±0.0186	±0.0184	±0.0179	±0.291
0.85	0.1498	0.6004	0.6022	0.7252	0.7496	0.8968	2.281
	±0.00217	±0.0121	±0.0133	±0.0122	±0.0126	±0.00757	±0.137

Table 23: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $H_2/M/s$ *model with* $s = 10$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.007826	0.03816	0.03816	0.03923	0.03970	0.04103	1.020
0.95	0.02101	0.09735	0.09734	0.1045	0.1077	0.1160	0.9996
0.93	0.02933	0.1321	0.1322	0.1455	0.1501	0.1662	0.9875
0.9	0.04281	0.1849	0.1851	0.2112	0.2155	0.2503	0.9928
0.85	0.06653	0.2666	0.2674	0.3220	0.3328	0.3982	1.0127

Table 24: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

ASE *in the* $H_2/M/s$ *model with* $s = 100$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.238×10^{-2}	6.040×10^{-2}	6.041×10^{-2}	6.612×10^{-2}	6.702×10^{-2}	1.034×10^{-1}	1.505
	6.95×10^{-4}	3.21×10^{-3}	3.22×10^{-3}	3.22×10^{-3}	3.22×10^{-3}	3.40×10^{-2}	0.2264
0.95	4.815×10^{-2}	2.248×10^{-2}	2.248×10^{-2}	2.765×10^{-2}	2.842×10^{-2}	5.633×10^{-2}	2.433×10^{-1}
	9.469×10^{-5}	4.628×10^{-4}	4.716×10^{-4}	4.527×10^{-4}	4.489×10^{-4}	5.797×10^{-4}	2.274×10^{-2}
0.93	3.444×10^{-3}	1.553×10^{-2}	1.554×10^{-2}	2.039×10^{-2}	2.108×10^{-2}	4.454×10^{-2}	1.214×10^{-1}
	9.45×10^{-5}	4.38×10^{-4}	4.44×10^{-4}	4.91×10^{-4}	5.00×10^{-4}	1.02×10^{-3}	1.02×10^{-2}
0.9	2.350×10^{-3}	1.020×10^{-2}	1.021×10^{-2}	1.460×10^{-2}	1.519×10^{-2}	3.305×10^{-2}	5.536×10^{-2}
	4.03×10^{-5}	2.11×10^{-4}	2.12×10^{-4}	2.36×10^{-4}	2.44×10^{-4}	5.28×10^{-4}	2.92×10^{-3}

Table 25: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ in the $H_2/M/s$ model with $s = 100$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.007519	0.03670	0.03670	0.04017	0.04072	0.06281	0.9146
0.95	0.02057	0.09600	0.09603	0.1181	0.1214	0.2406	1.039
0.93	0.02927	0.1320	0.1321	0.1733	0.1792	0.3786	1.031
0.9	0.04213	0.1829	0.1830	0.2618	0.2723	0.5924	0.9906

Table 26: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

			ASE in the $H_2/M/s$ *model with* $s = 400$				
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	7.676×10^{-4}	3.736×10^{-3}	3.736×10^{-3}	4.449×10^{-3}	4.561×10^{-3}	1.329×10^{-2}	9.183×10^{-2}
	2.34×10^{-5}	1.05×10^{-4}	1.05×10^{-4}	1.04×10^{-4}	1.05×10^{-4}	1.64×10^{-4}	6.34×10^{-3}
0.95	3.018×10^{-4}	1.382×10^{-3}	1.382×10^{-3}	2.015×10^{-3}	2.105×10^{-3}	7.525×10^{-3}	1.423×10^{-2}
	8.57×10^{-6}	4.60×10^{-5}	4.67×10^{-5}	4.99×10^{-5}	5.14×10^{-5}	1.406×10^{-4}	1.471×10^{-3}
0.93	2.16×10^{-4}	9.74×10^{-4}	9.752×10^{-4}	1.561×10^{-3}	1.635×10^{-3}	5.736×10^{-3}	7.613×10^{-3}
	1.07×10^{-5}	3.70×10^{-5}	3.79×10^{-5}	4.12×10^{-5}	4.14×10^{-5}	1.93×10^{-4}	1.157×10^{-3}
0.9	1.443×10^{-4}	6.294×10^{-4}	6.300×10^{-4}	1.133×10^{-3}	1.194×10^{-3}	3.798×10^{-3}	3.368×10^{-3}
	4.09×10^{-6}	2.56×10^{-5}	2.59×10^{-5}	3.40×10^{-5}	3.75×10^{-5}	1.55×10^{-4}	2.32×10^{-4}

Table 27: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $H_2/M/s$ *model with* $s = 400$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.008655	0.04212	0.04212	0.05016	0.05142	0.1498	1.035
0.95	0.02034	0.09315	0.09315	0.1358	0.1419	0.5072	0.9590
0.93	0.02909	0.1314	0.1315	0.2098	0.2205	0.7735	1.027
0.9	0.04224	0.1842	0.1844	0.3316	0.3495	1.112	0.9858

Table 28: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

			ASE in the $H_2/M/s$ model with $s = 900$				
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.487×10^{-4}	7.206×10^{-4}	7.205×10^{-4}	9.297×10^{-4}	9.622×10^{-4}	4.395×10^{-3}	1.698×10^{-2}
	7.56×10^{-6}	3.42×10^{-5}	3.45×10^{-5}	3.36×10^{-5}	3.38×10^{-5}	8.20×10^{-5}	2.156×10^{-3}
0.95	5.826×10^{-5}	2.713×10^{-4}	2.713×10^{-4}	4.512×10^{-4}	4.757×10^{-4}	2.276×10^{-3}	2.539×10^{-3}
	3.70×10^{-6}	1.72×10^{-5}	1.75×10^{-5}	2.15×10^{-5}	2.21×10^{-5}	1.85×10^{-4}	3.354×10^{-4}
0.93	4.292×10^{-5}	1.935×10^{-4}	1.936×10^{-4}	3.619×10^{-4}	3.832×10^{-4}	1.630×10^{-3}	1.440×10^{-3}
	2.12×10^{-6}	1.29×10^{-5}	1.30×10^{-5}	1.68×10^{-5}	1.75×10^{-5}	1.32×10^{-4}	2.37×10^{-4}

Table 29: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval.

$RASE$ *in the* $H_2/M/s$ *model with* $s = 900$							
ρ	$RASE(QL)$	$RASE(LES)$	$RASE(HOL)$	$RASE(RCS)$	$RASE(RCS-\sqrt{s})$	$RASE(LCS)$	$RASE(NI)$
0.98	0.007676	0.03720	0.03720	0.04799	0.04967	0.2268	0.8767
0.95	0.02056	0.09582	0.09579	0.15932	0.1680	0.8039	0.8965
0.93	0.02956	0.1332	0.1333	0.2492	0.2638	1.122	0.9916

Table 30: A comparison of the efficiency of different real-time delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the relative average squared error – (RASE).

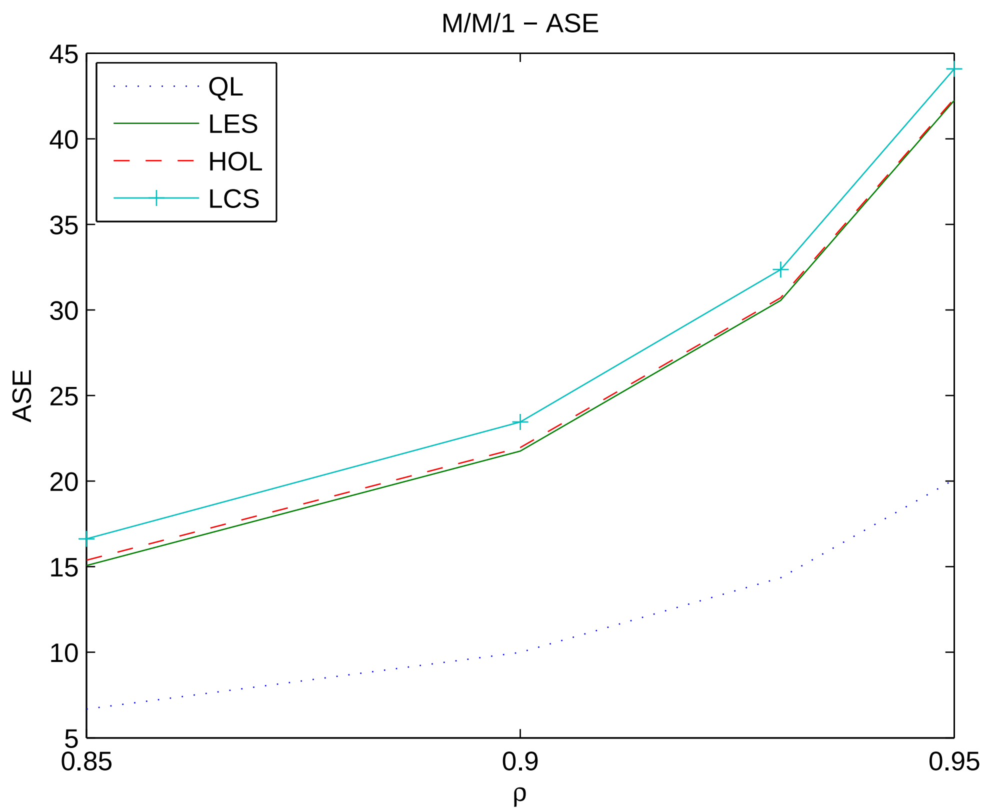

Figure 1: Point estimates of the ASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

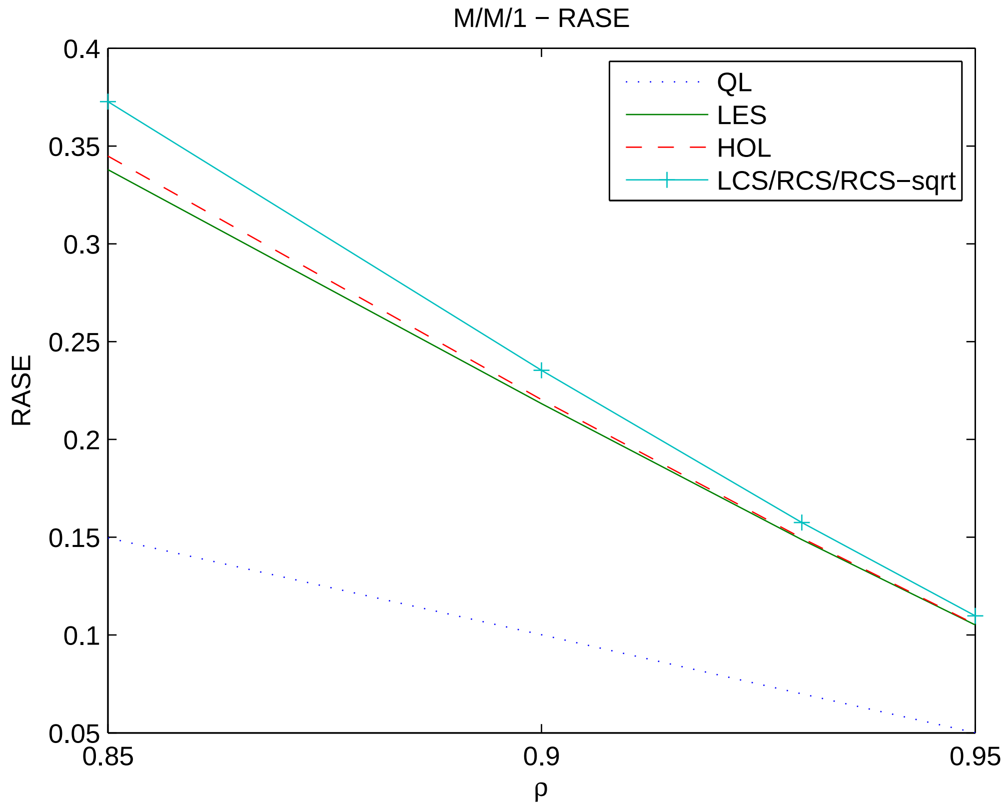

Figure 2: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

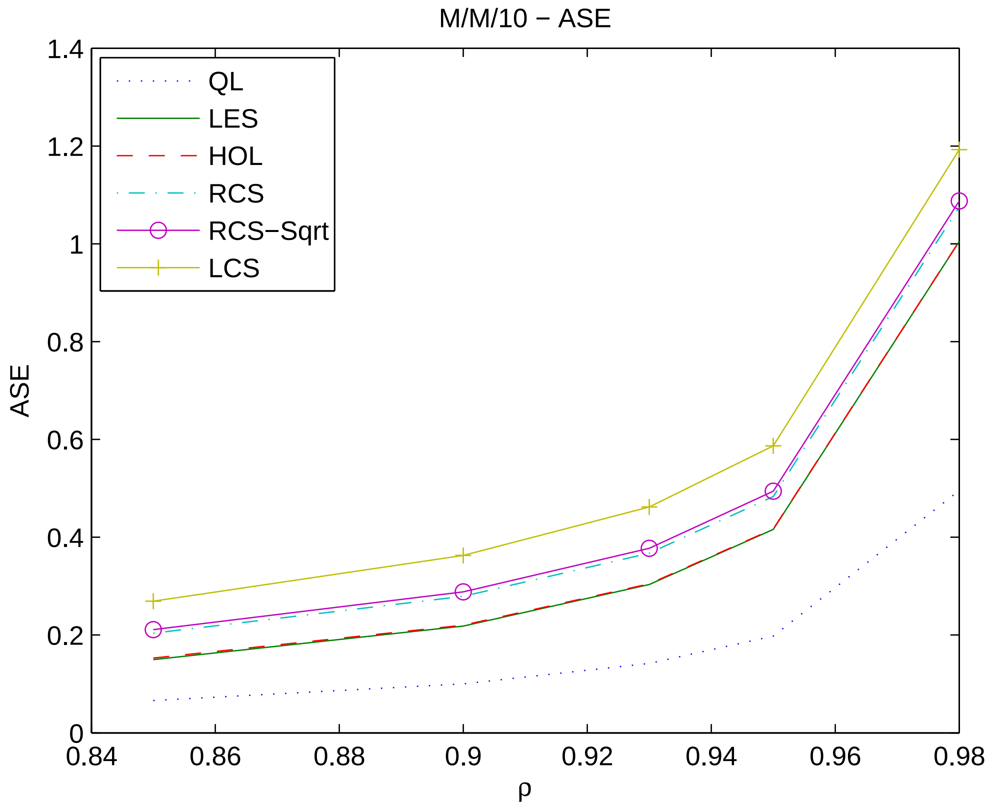

Figure 3: Point estimates of the ASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

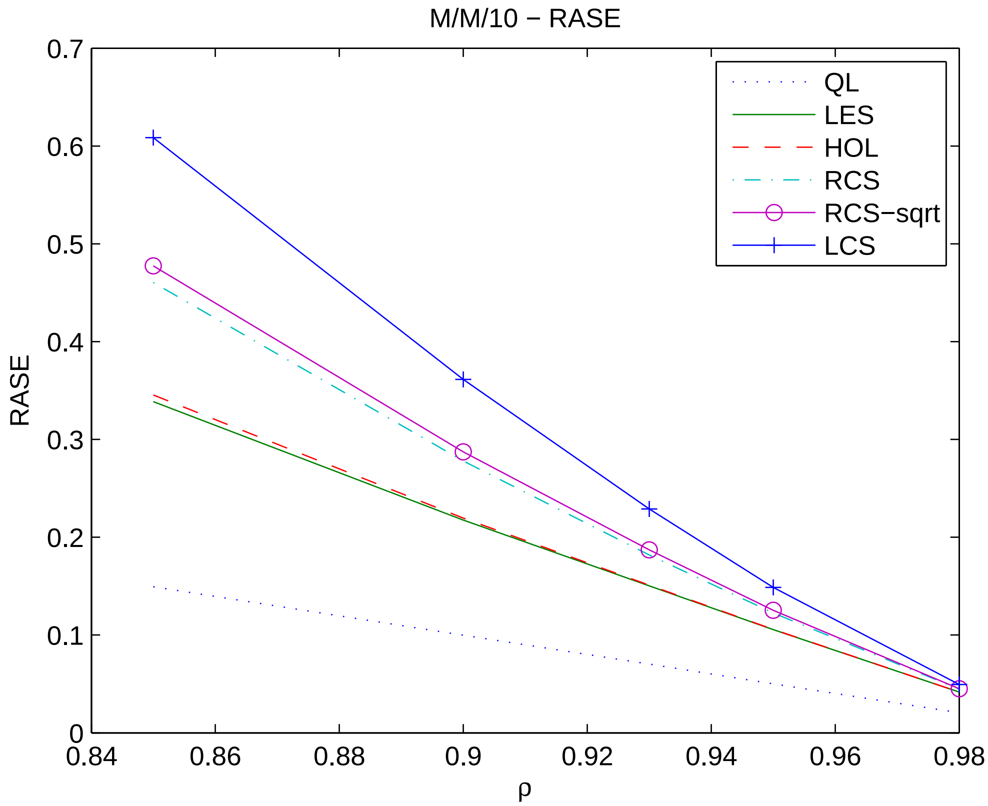

Figure 4: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

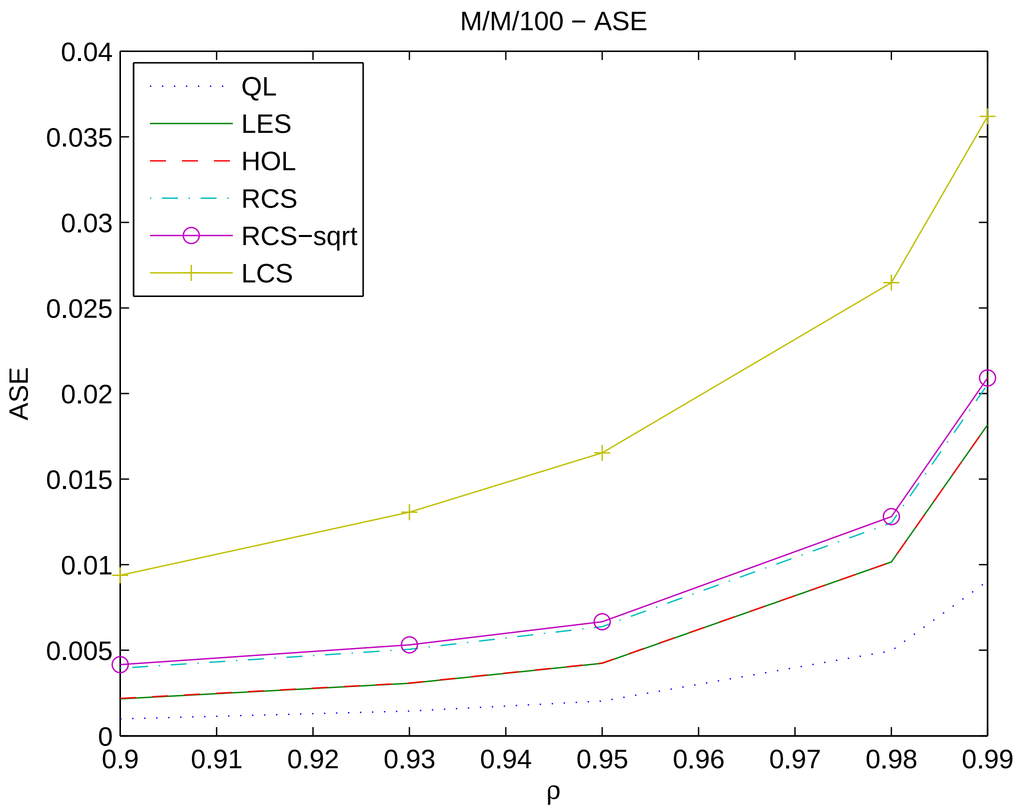

Figure 5: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

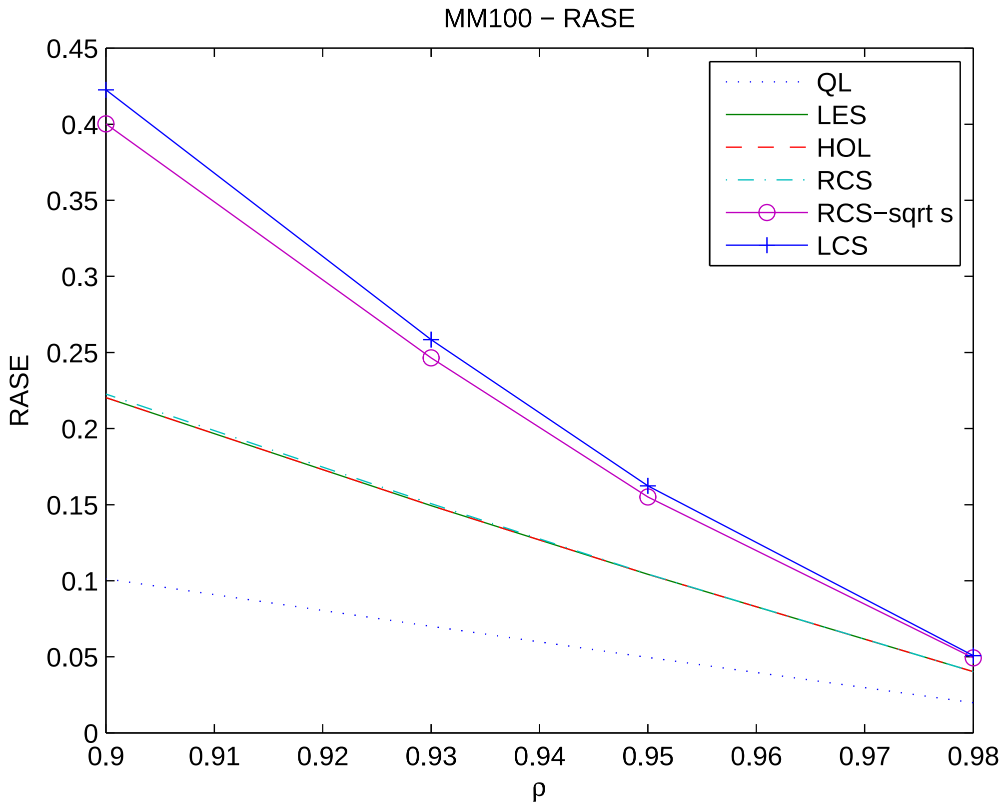

Figure 6: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

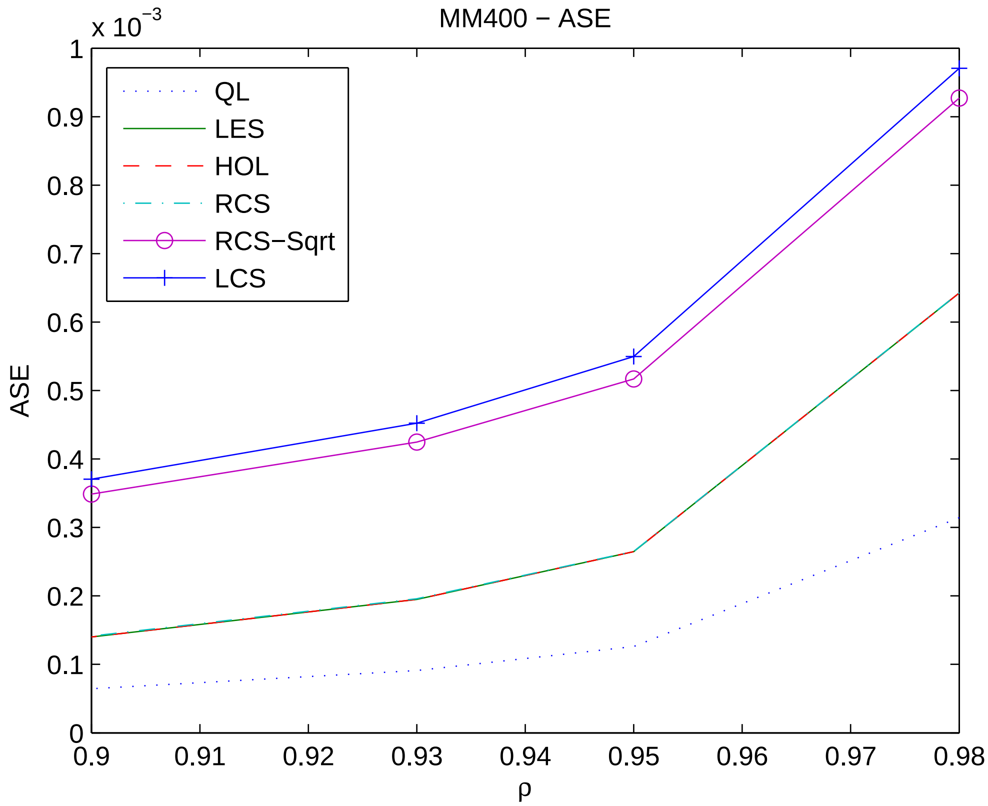

Figure 7: Point estimates of the ASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

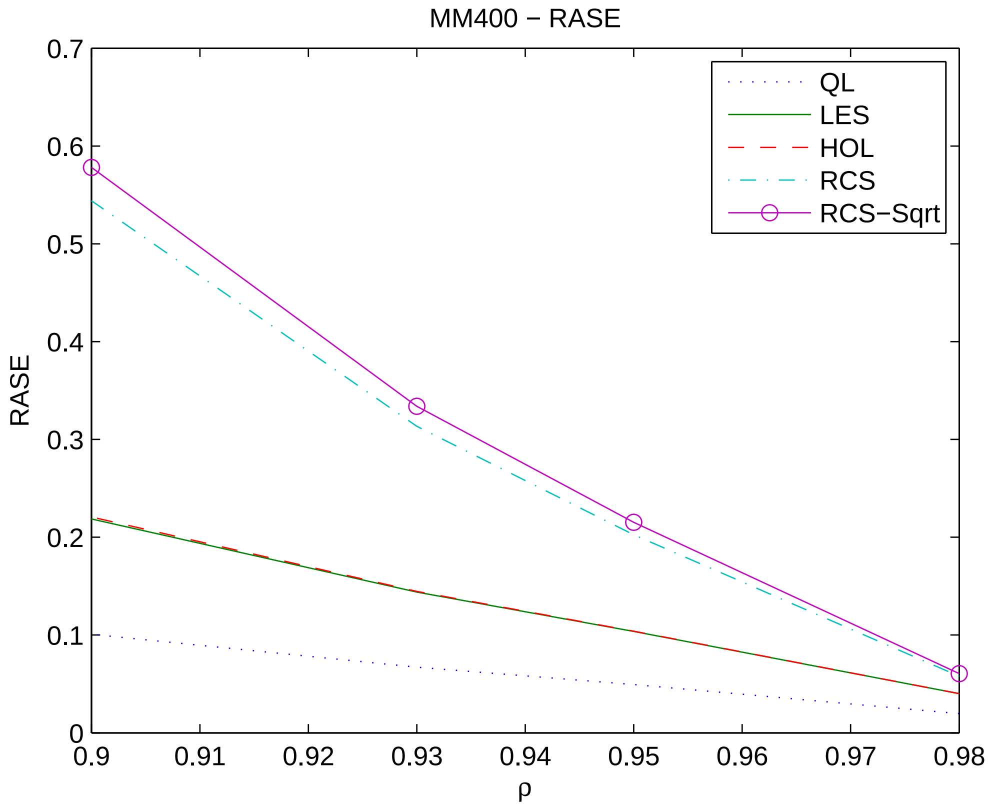

Figure 8: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

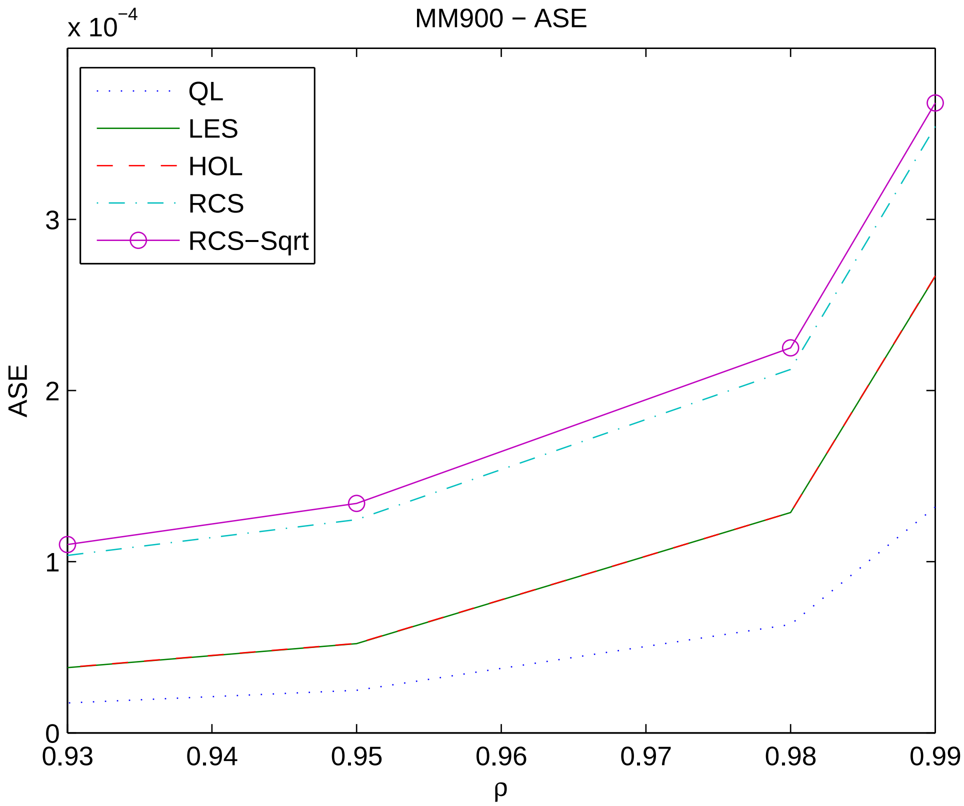

Figure 9: Point estimates of the ASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

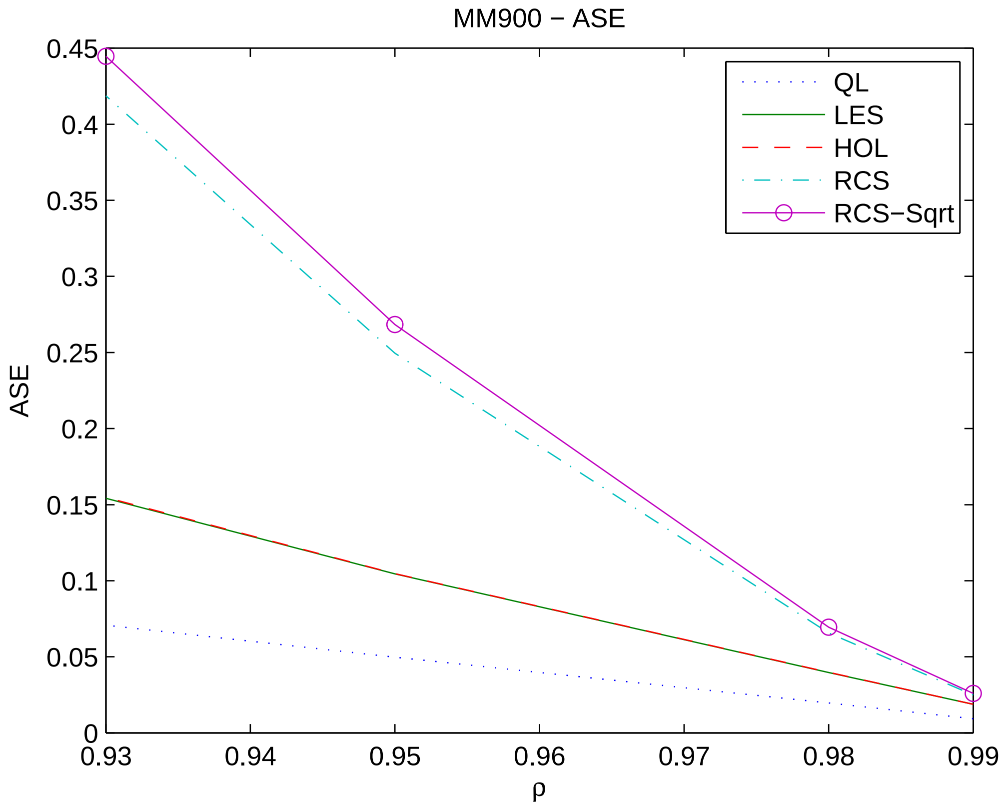

Figure 10: Point estimates of the RASE of alternative real-time delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

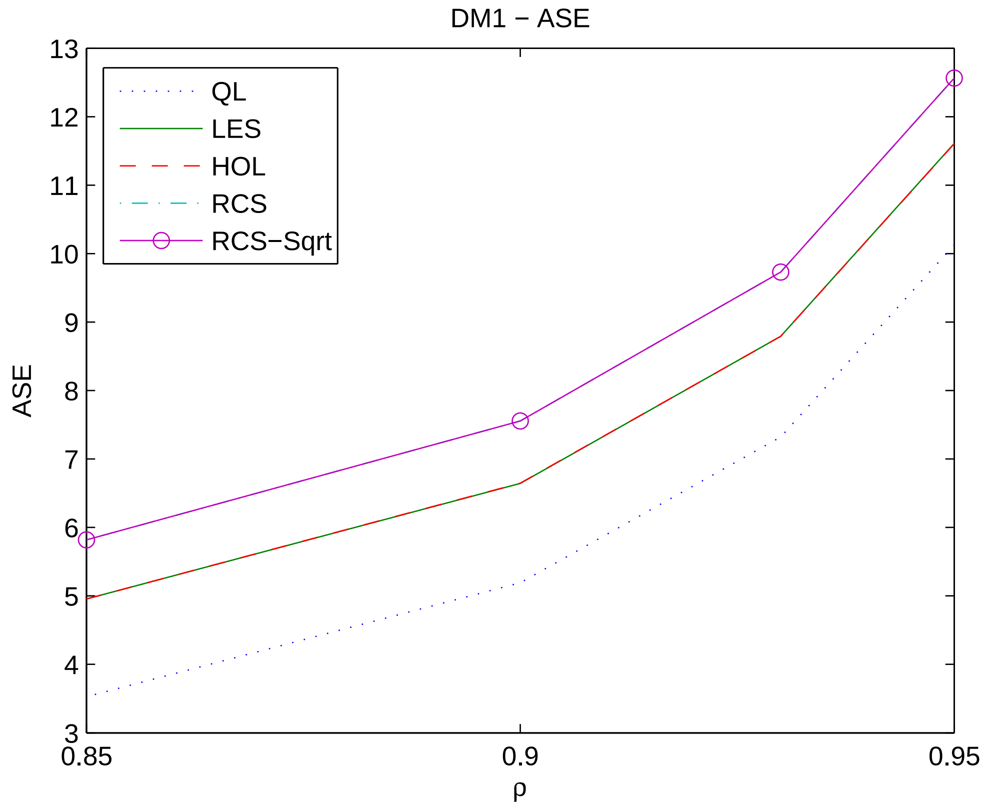

Figure 11: Point estimates of the ASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

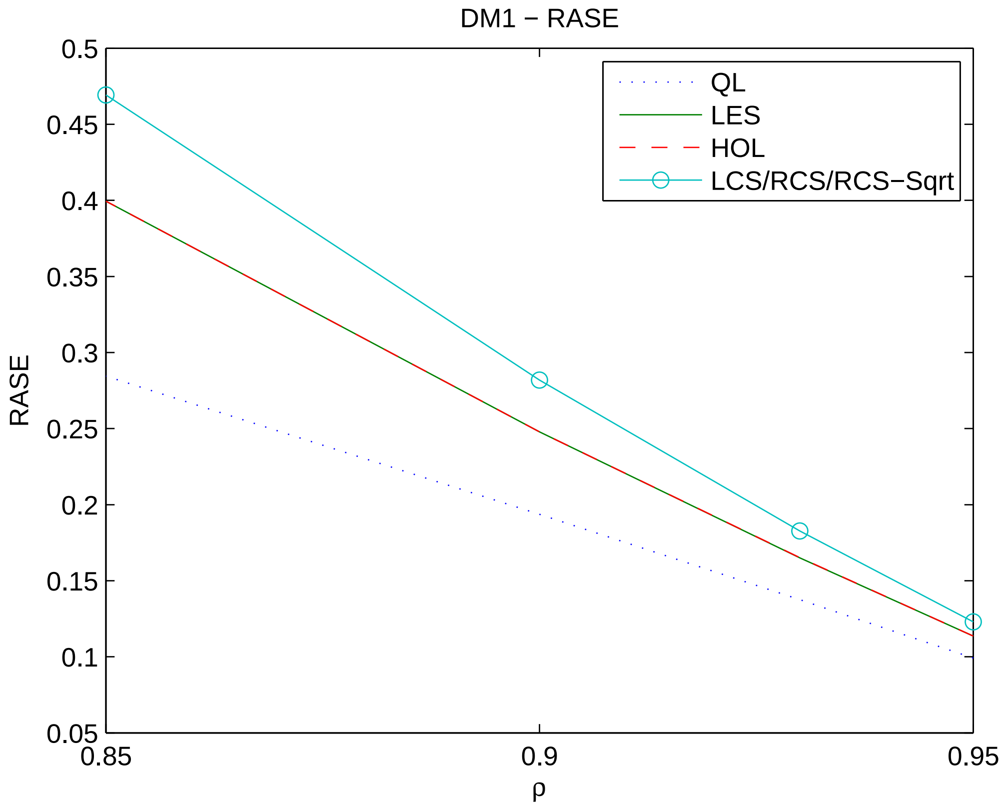

Figure 12: Point estimates of the RASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

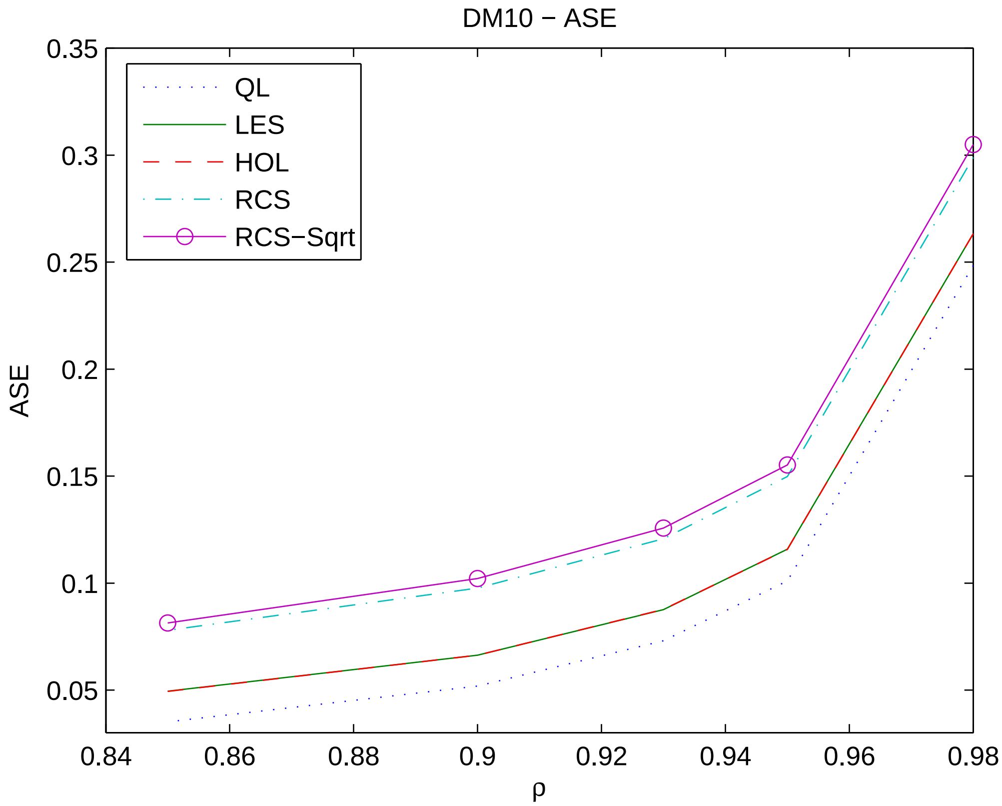

Figure 13: Point estimates of the ASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

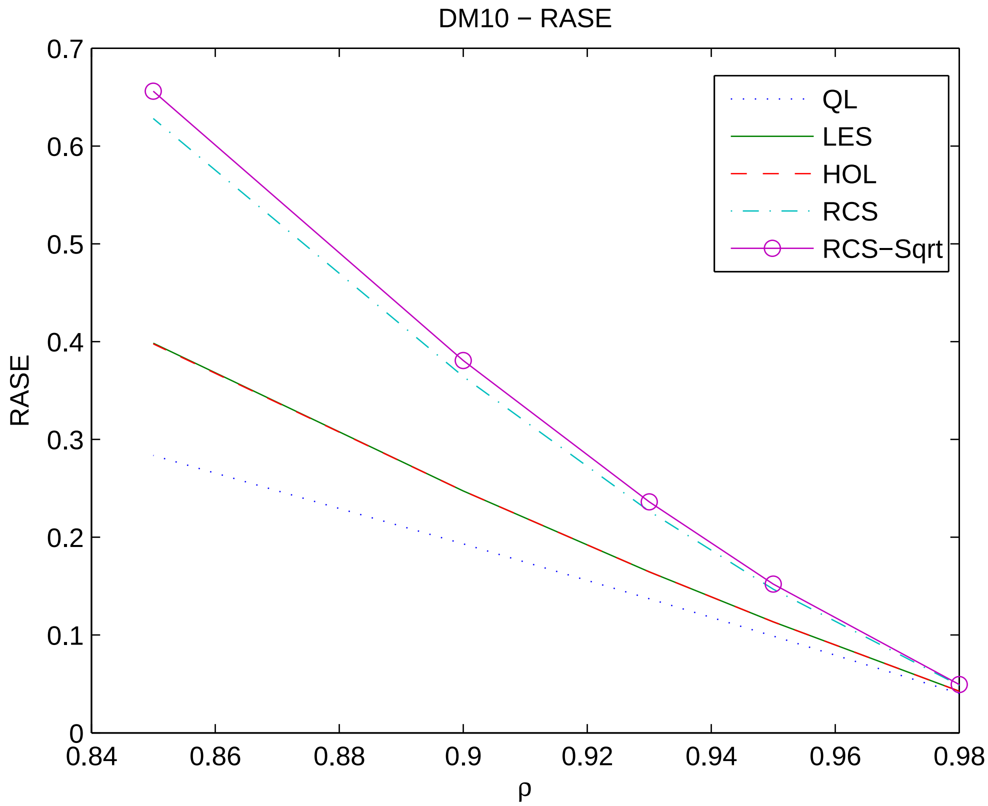

Figure 14: Point estimates of the RASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

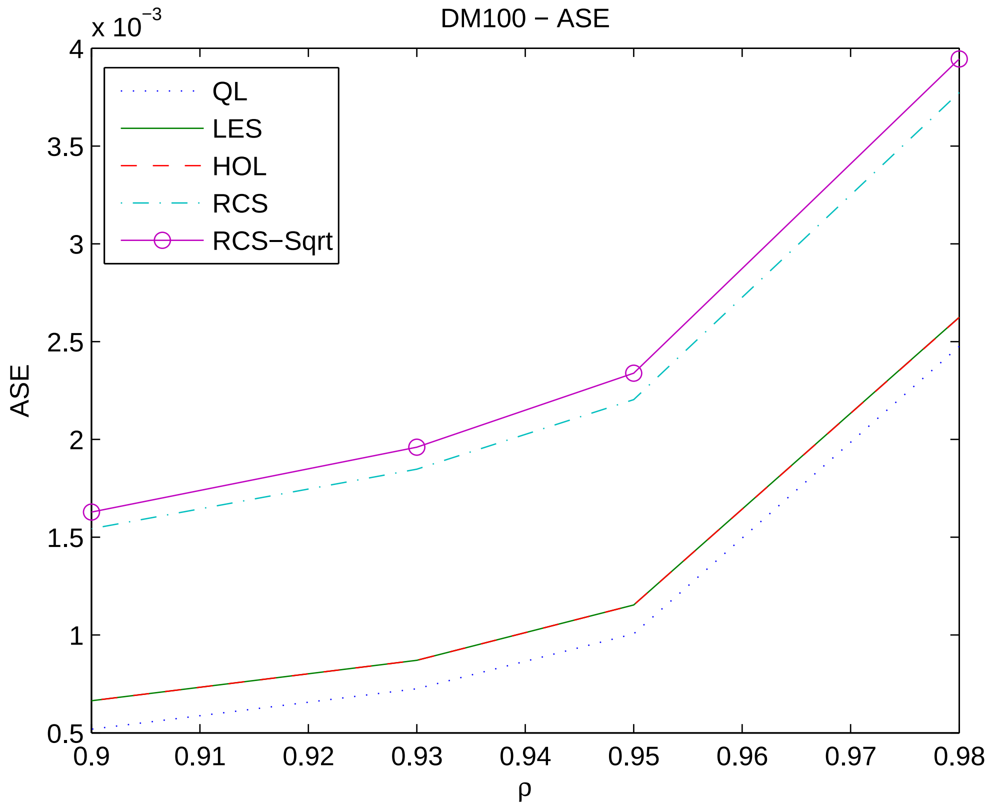

Figure 15: Point estimates of the ASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

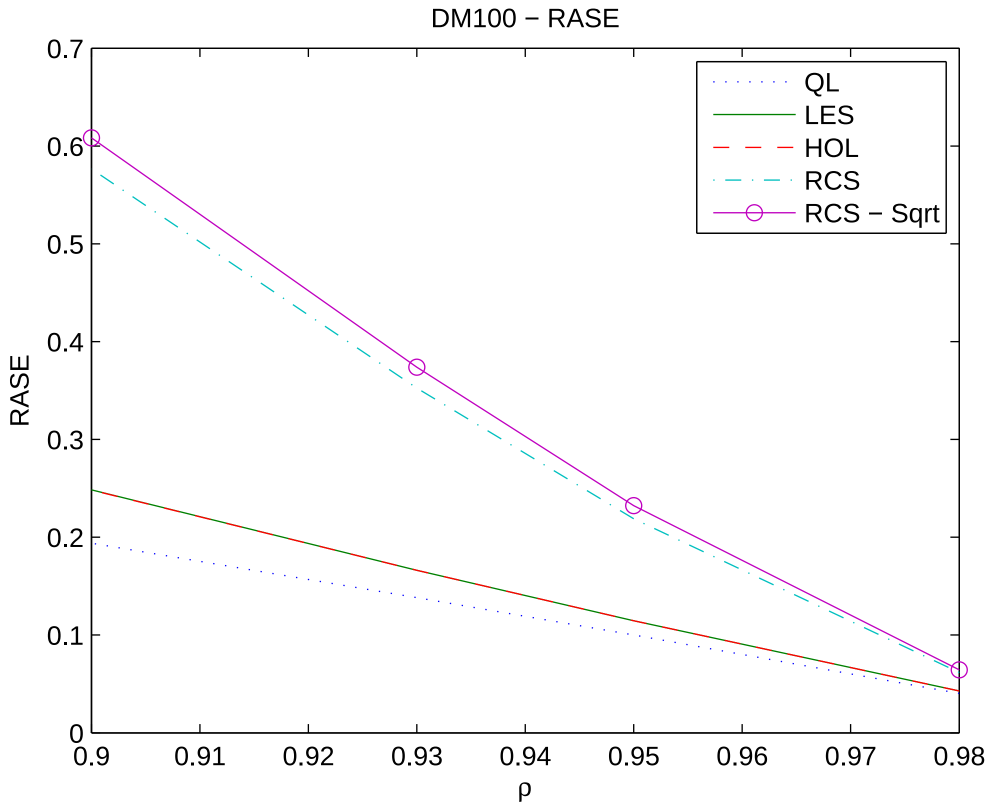

Figure 16: Point estimates of the RASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

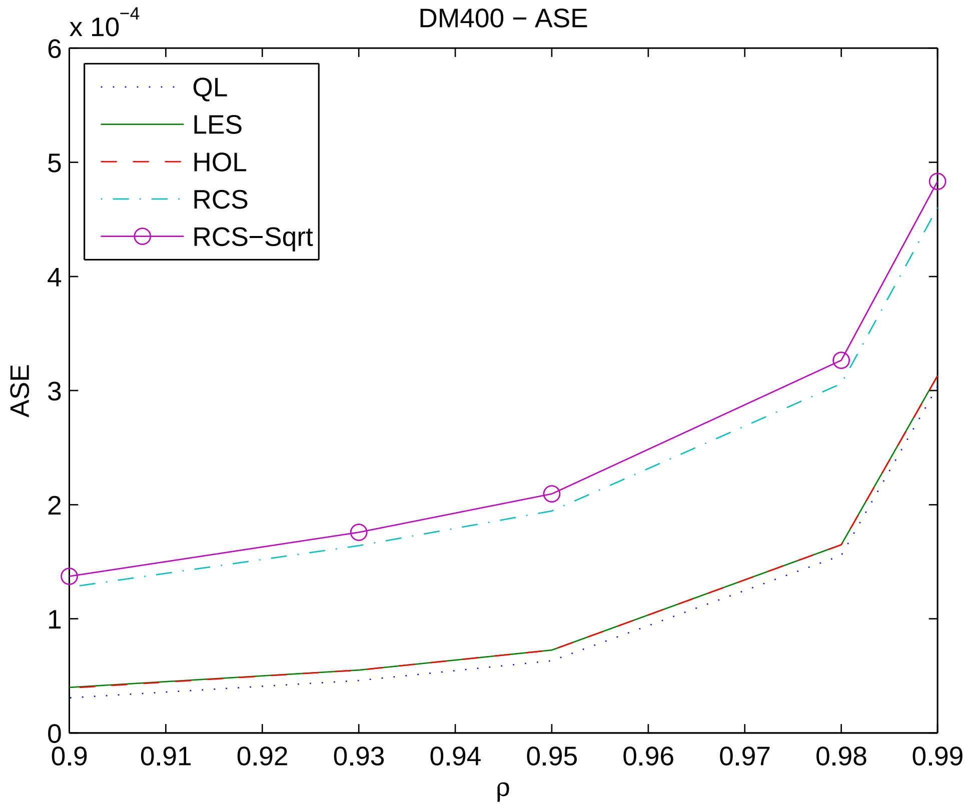

Figure 17: Point estimates of the ASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

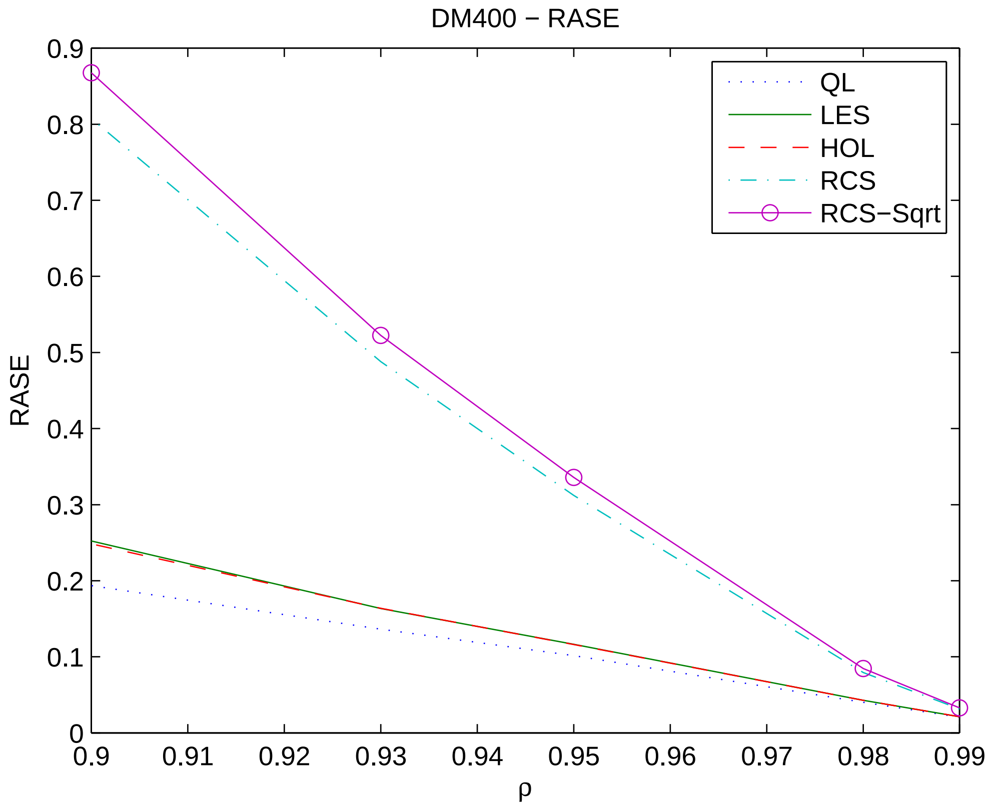

Figure 18: Point estimates of the RASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

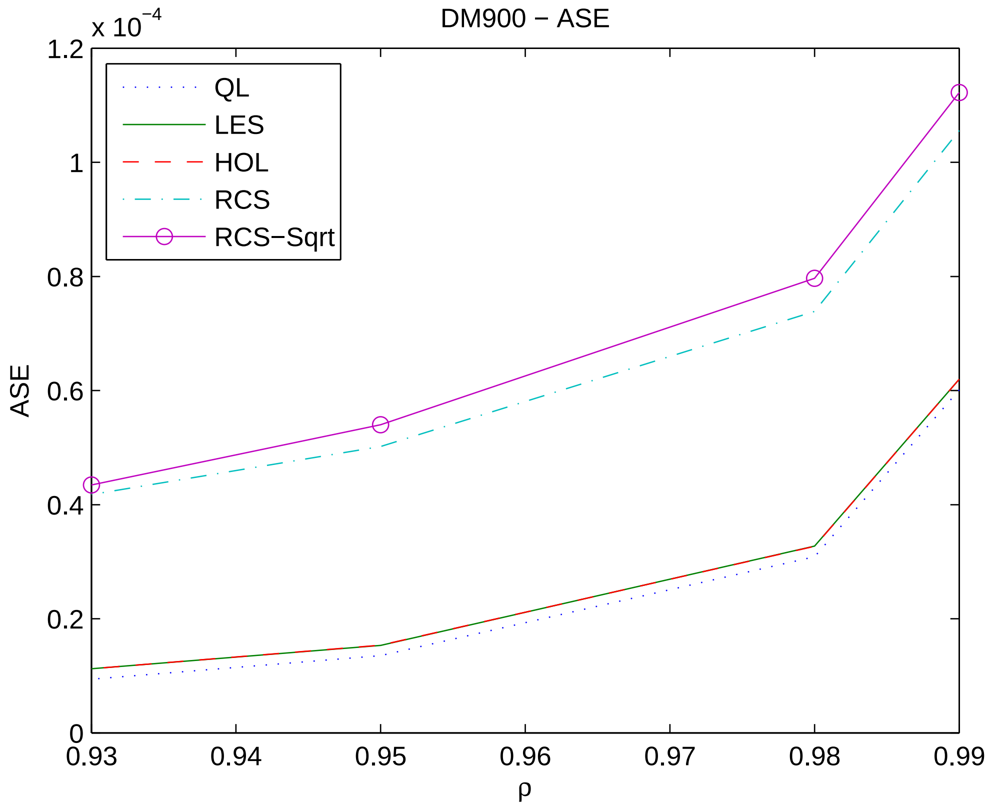

Figure 19: Point estimates of the ASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

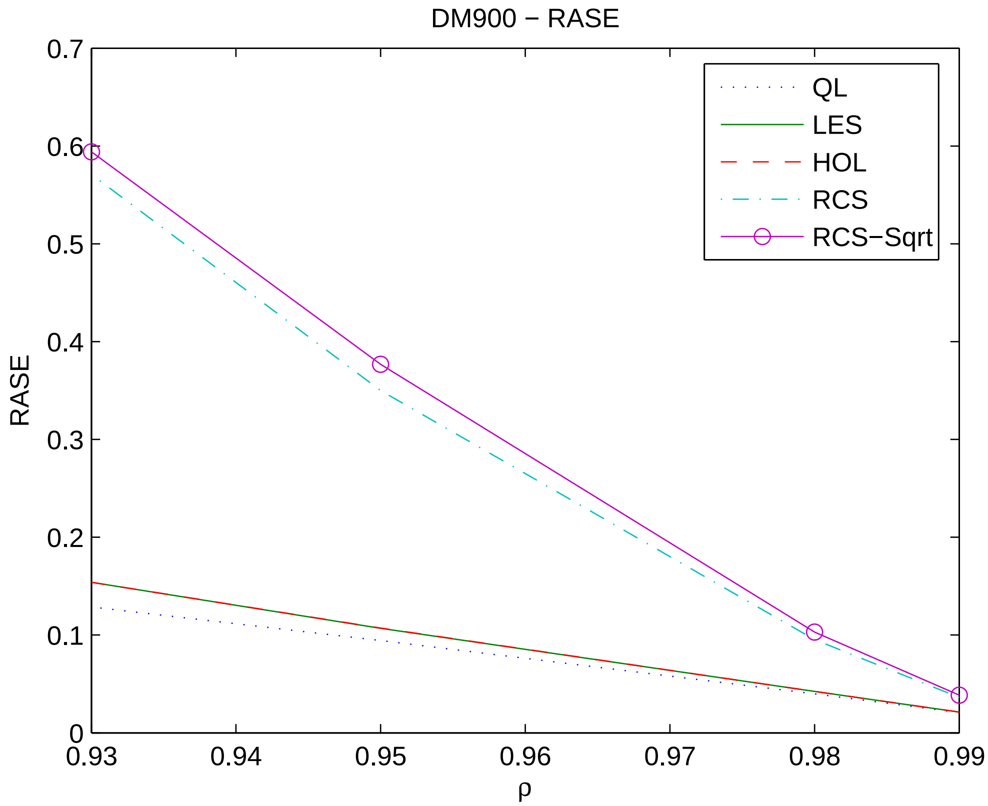

Figure 20: Point estimates of the RASE of alternative real-time delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

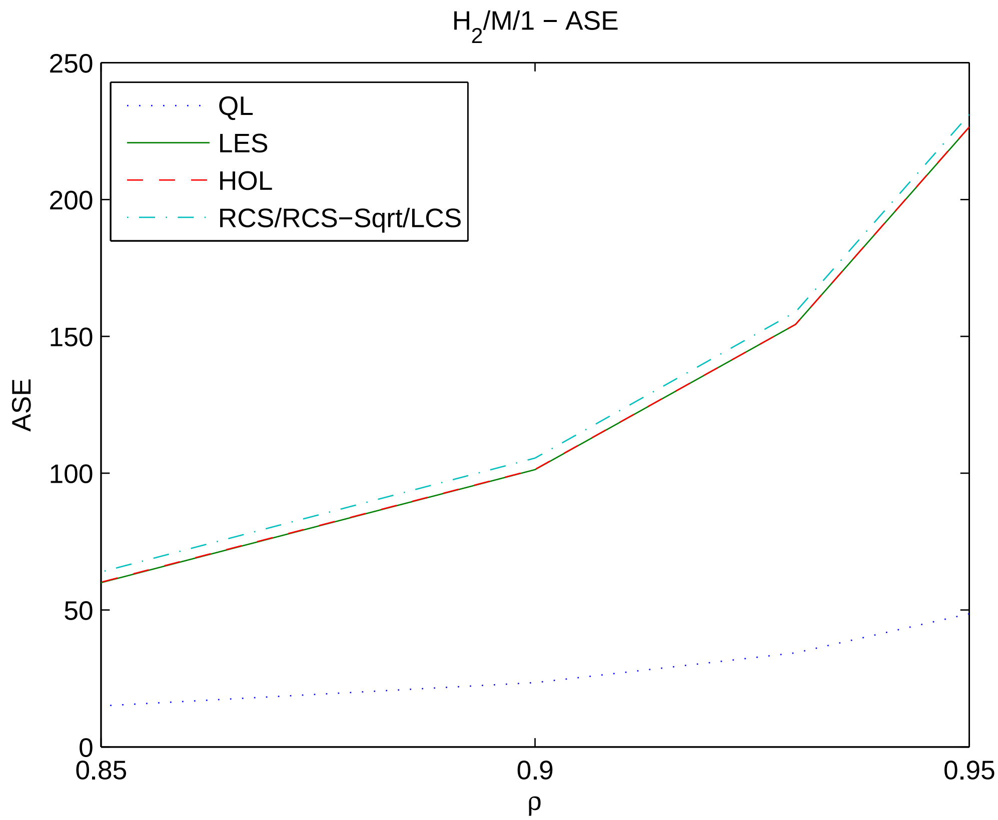

Figure 21: Point estimates of the ASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

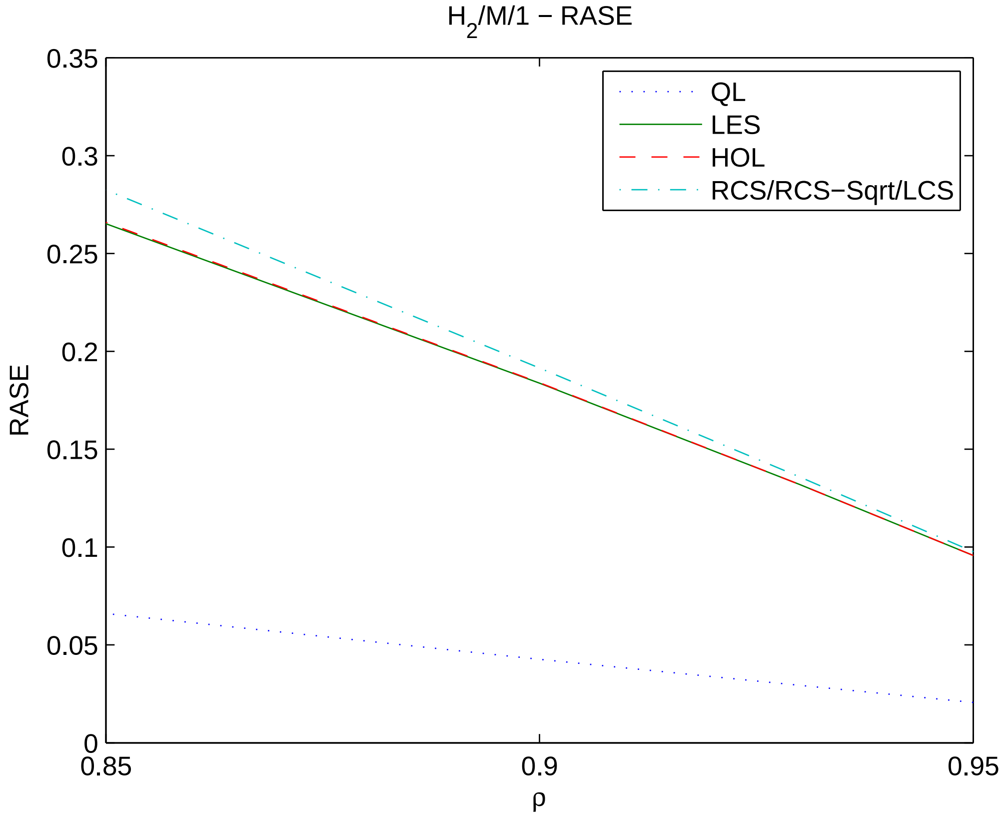

Figure 22: Point estimates of the RASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ.

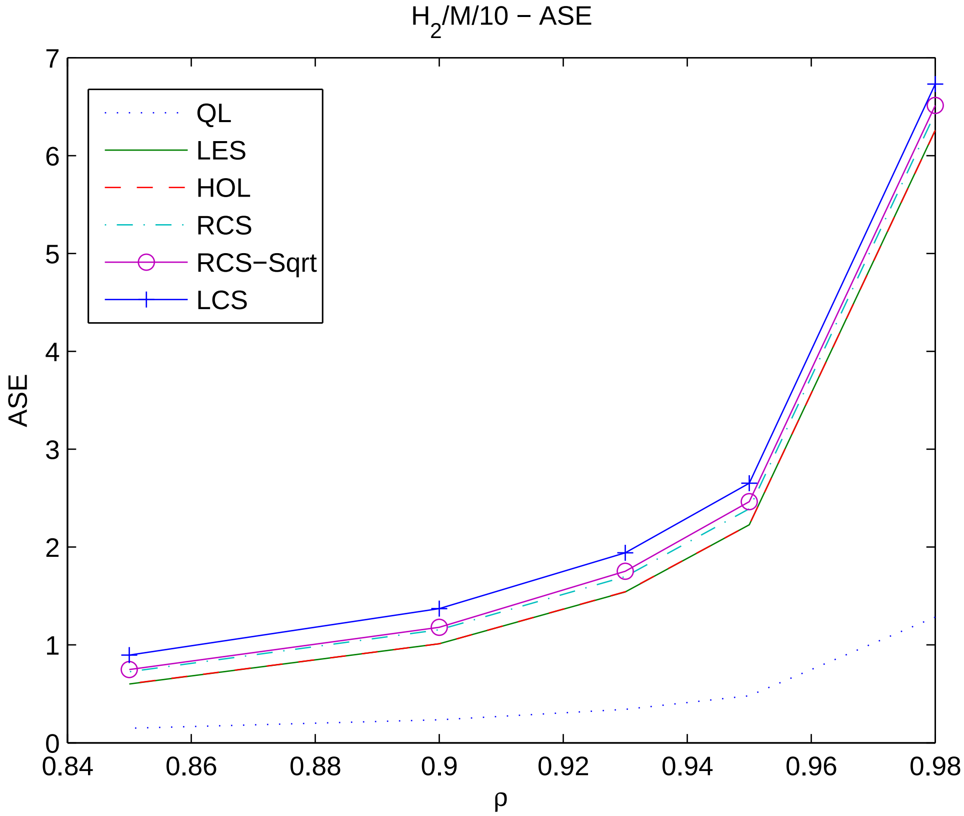

Figure 23: Point estimates of the ASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

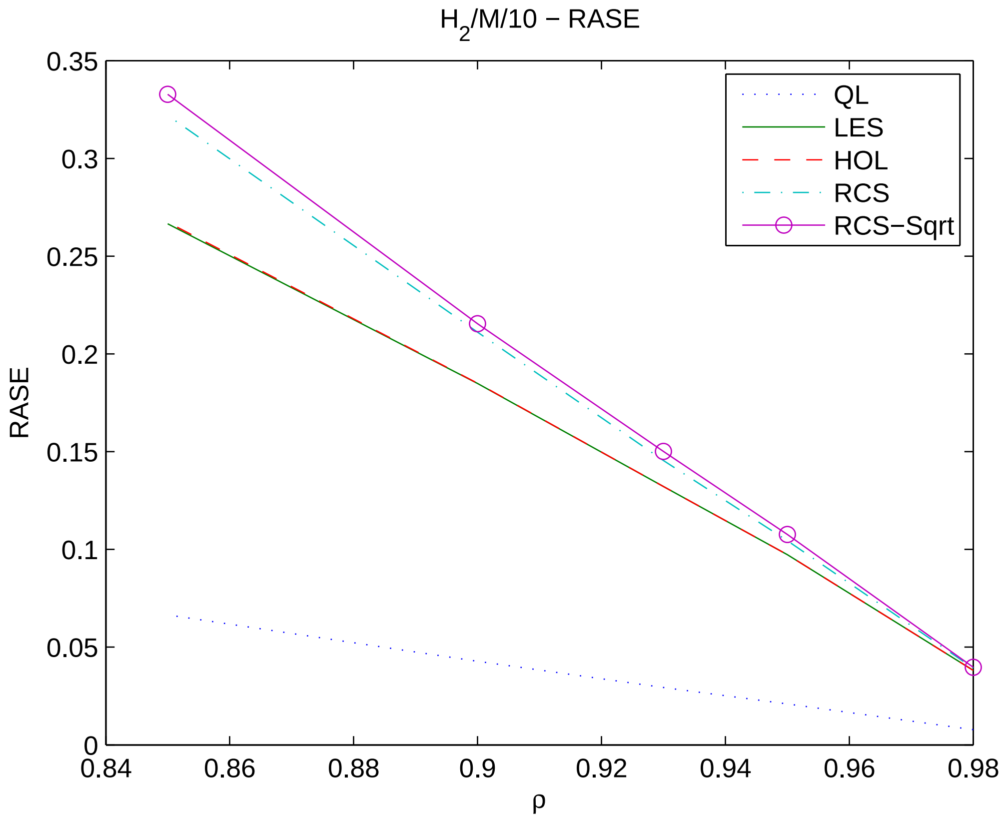

Figure 24: Point estimates of the RASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ.

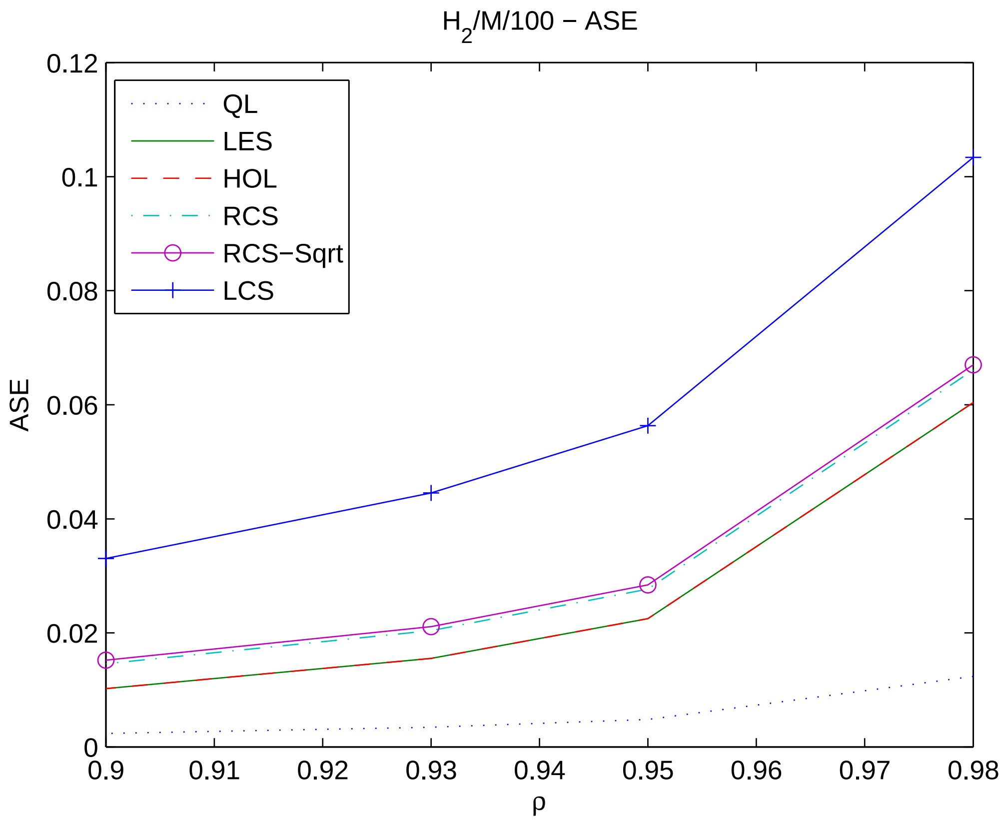

Figure 25: Point estimates of the ASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

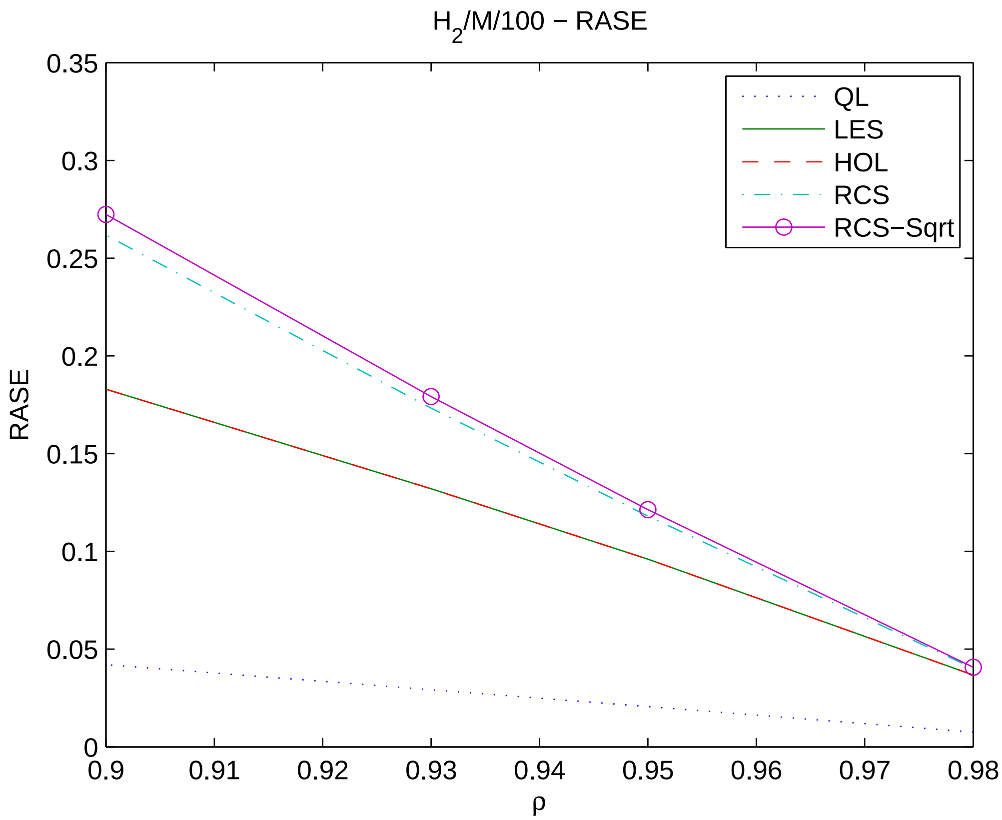

Figure 26: Point estimates of the RASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ.

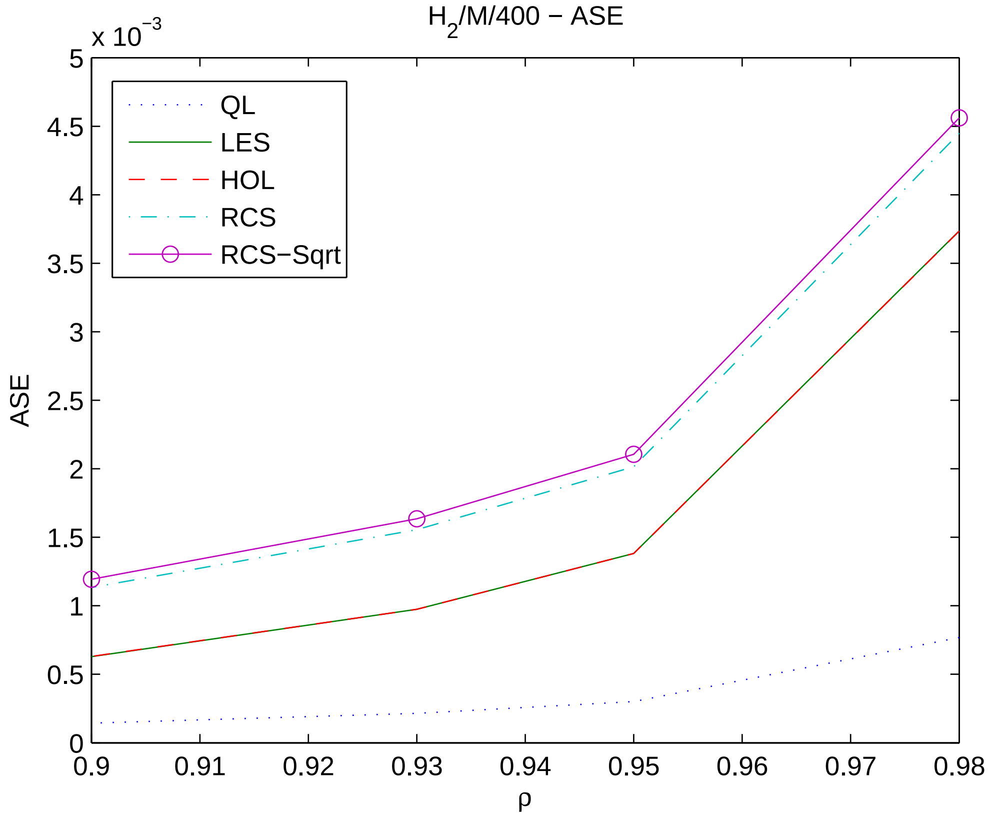

Figure 27: Point estimates of the ASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

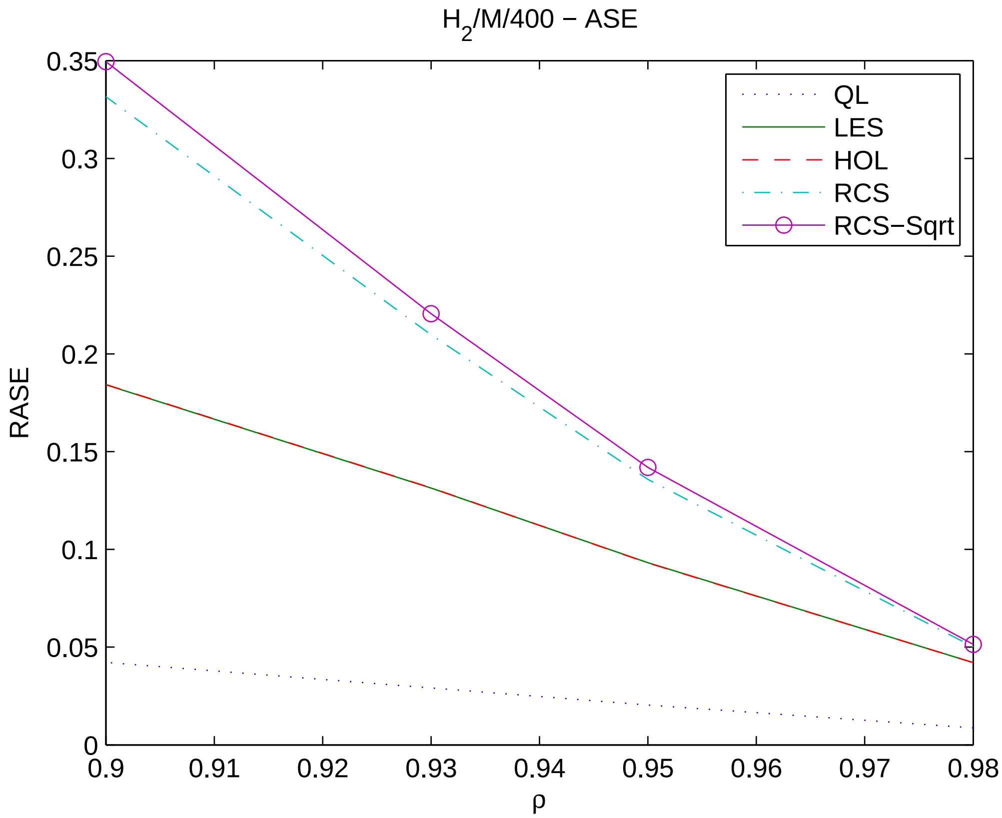

Figure 28: Point estimates of the RASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ.

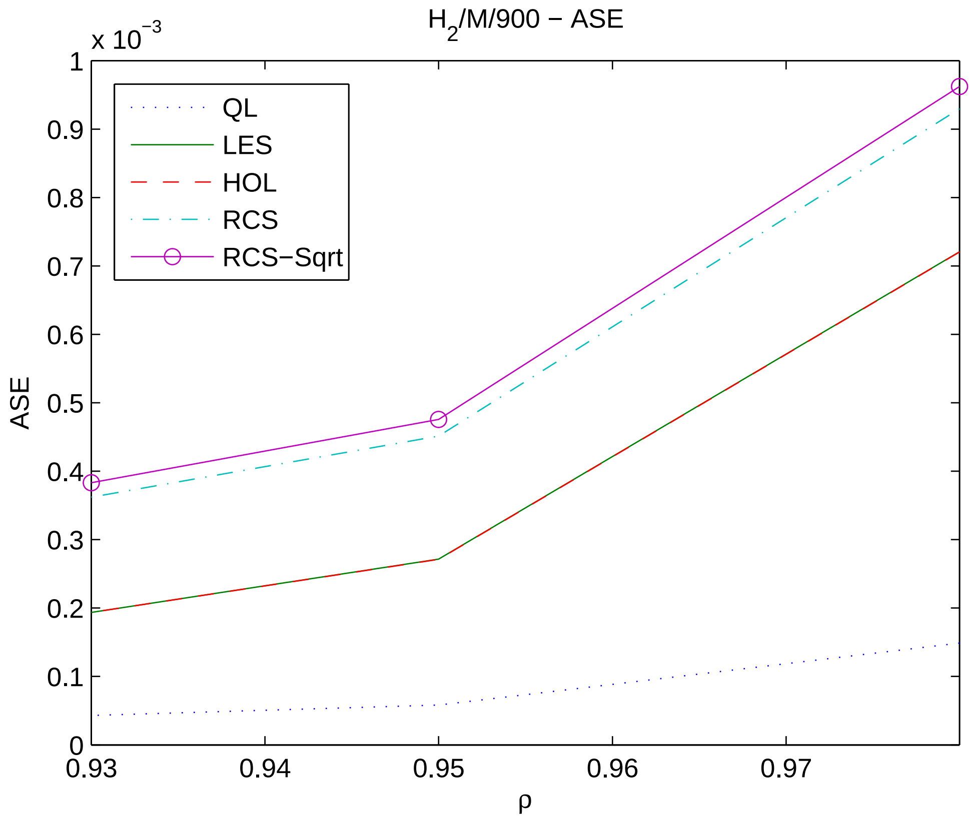

Figure 29: Point estimates of the ASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

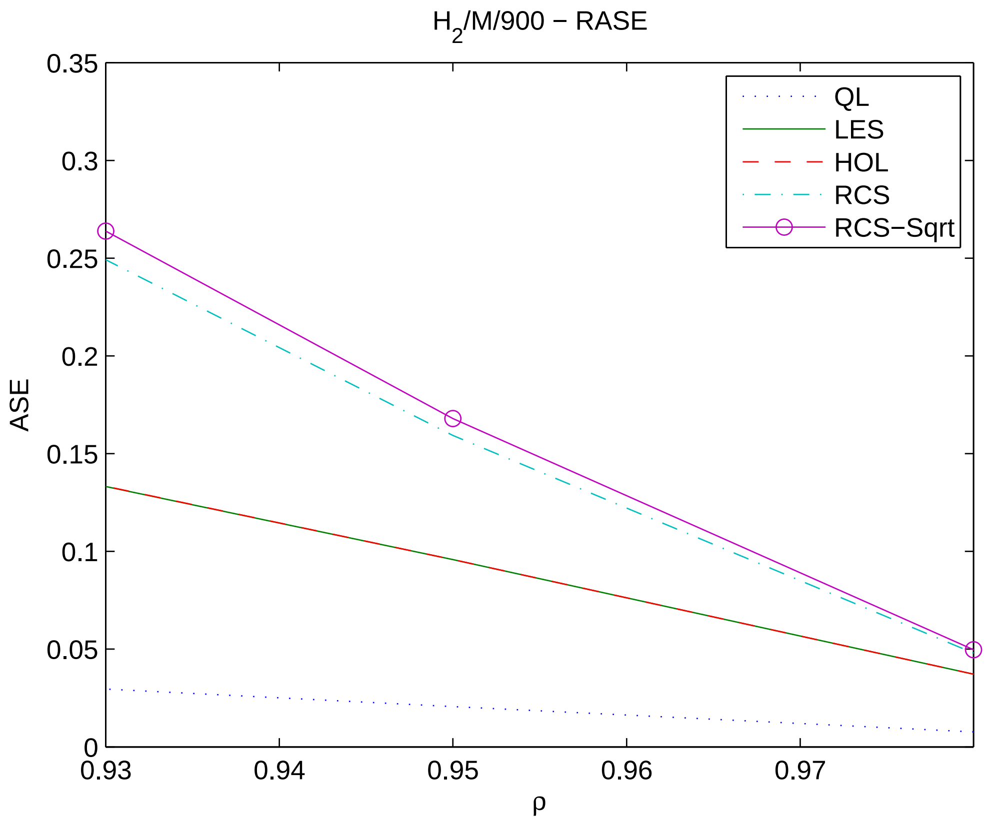

Figure 30: Point estimates of the RASE of alternative real-time delay estimators for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ.

Conditional ASE in the $M/M/s$ model with $s = 100$ for actual delays in $(\widehat{E[W\|W > 0]}, 2\widehat{E[W\|W > 0]})$								
ρ	$\widehat{E[W\|W > 0]}$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	9.833×10^{-1}	1.380×10^{-2}	2.802×10^{-2}	2.782×10^{-2}	3.052×10^{-2}	3.090×10^{-2}	4.884×10^{-2}	2.504×10^{-2}
	$\pm 6.92 \times 10^{-2}$	$\pm 9.71 \times 10^{-4}$	$\pm 2.08 \times 10^{-3}$	$\pm 2.08 \times 10^{-3}$	$\pm 2.09 \times 10^{-3}$	$\pm 2.09 \times 10^{-3}$	$\pm 2.18 \times 10^{-3}$	$\pm 3.89 \times 10^{-2}$
0.98	5.039×10^{-1}	7.006×10^{-3}	1.442×10^{-2}	1.421×10^{-2}	1.702×10^{-2}	1.744×10^{-2}	3.565×10^{-2}	6.453×10^{-2}
	$\pm 2.48 \times 10^{-2}$	$\pm 3.47 \times 10^{-4}$	$\pm 7.41 \times 10^{-4}$	$\pm 7.42 \times 10^{-4}$	$\pm 7.62 \times 10^{-4}$	$\pm 7.62 \times 10^{-4}$	$\pm 9.67 \times 10^{-4}$	$\pm 6.08 \times 10^{-3}$
0.95	2.028×10^{-1}	2.772×10^{-3}	5.924×10^{-3}	5.695×10^{-3}	8.711×10^{-3}	9.145×10^{-3}	2.339×10^{-2}	1.040×10^{-2}
	$\pm 3.02 \times 10^{-3}$	$\pm 5.46 \times 10^{-5}$	$\pm 1.07 \times 10^{-4}$	$\pm 1.06 \times 10^{-4}$	$\pm 1.37 \times 10^{-4}$	$\pm 1.44 \times 10^{-4}$	$\pm 4.23 \times 10^{-4}$	$\pm 2.58 \times 10^{-4}$
0.93	1.435×10^{-1}	1.925×10^{-3}	4.231×10^{-3}	3.996×10^{-3}	7.031×10^{-3}	7.421×10^{-3}	1.776×10^{-2}	5.263×10^{-3}
	$\pm 1.78 \times 10^{-3}$	$\pm 2.72 \times 10^{-5}$	$\pm 6.17 \times 10^{-5}$	$\pm 6.00 \times 10^{-5}$	$\pm 8.37 \times 10^{-5}$	$\pm 8.77 \times 10^{-5}$	$\pm 2.67 \times 10^{-4}$	$\pm 1.20 \times 10^{-4}$
0.90	9.929×10^{-2}	1.293×10^{-3}	2.982×10^{-3}	2.734×10^{-3}	5.564×10^{-3}	5.850×10^{-3}	1.202×10^{-2}	2.516×10^{-3}
	$\pm 2.75 \times 10^{-3}$	$\pm 4.18 \times 10^{-5}$	$\pm 8.61 \times 10^{-5}$	$\pm 8.47 \times 10^{-5}$	$\pm 1.44 \times 10^{-4}$	$\pm 1.55 \times 10^{-4}$	$\pm 4.45 \times 10^{-4}$	$\pm 1.27 \times 10^{-4}$

Table 31: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(\widehat{E[W|W > 0]}, 2\widehat{E[W|W > 0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the M/M/s model with $s = 100$ *for actual delays in* $(2E[\widehat{W\|W} > 0], 4E[\widehat{W\|W} > 0])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	2.589×10^{-2}	5.128×10^{-2}	5.108×10^{-2}	5.373×10^{-2}	5.417×10^{-2}	7.176×10^{-2}	2.882
	$\pm 1.56 \times 10^{-3}$	$\pm 2.77 \times 10^{-3}$	$\pm 2.77 \times 10^{-3}$	$\pm 2.79 \times 10^{-3}$	$\pm 2.79 \times 10^{-3}$	$\pm 2.93 \times 10^{-3}$	$\pm 3.81 \times 10^{-1}$
0.98	2.713×10^{-3}	5.842×10^{-3}	5.556×10^{-3}	9.434×10^{-3}	1.000×10^{-2}	2.731×10^{-2}	3.058×10^{-2}
	$\pm 7.57 \times 10^{-5}$	$\pm 1.57 \times 10^{-4}$	$\pm 1.58 \times 10^{-4}$	$\pm 1.89 \times 10^{-4}$	$\pm 1.92 \times 10^{-4}$	$\pm 6.12 \times 10^{-4}$	$\pm 1.62 \times 10^{-3}$
0.95	1.359×10^{-2}	2.727×10^{-2}	2.706×10^{-2}	2.986×10^{-2}	3.028×10^{-2}	4.974×10^{-2}	7.834×10^{-1}
	$\pm 6.75 \times 10^{-4}$	$\pm 1.35 \times 10^{-3}$	$\pm 1.35 \times 10^{-3}$	$\pm 1.27 \times 10^{-3}$	$\pm 1.26 \times 10^{-3}$	$\pm 1.04 \times 10^{-3}$	$\pm 9.44 \times 10^{-2}$
0.93	5.536×10^{-3}	1.141×10^{-2}	1.117×10^{-2}	1.445×10^{-2}	1.496×10^{-2}	3.762×10^{-2}	1.300×10^{-1}
	$\pm 1.26 \times 10^{-4}$	$\pm 1.47 \times 10^{-4}$	$\pm 1.46 \times 10^{-4}$	$\pm 1.56 \times 10^{-4}$	$\pm 1.79 \times 10^{-4}$	$\pm 6.06 \times 10^{-4}$	$\pm 4.04 \times 10^{-3}$
0.90	3.886×10^{-2}	8.119×10^{-3}	7.867×10^{-3}	1.133×10^{-2}	1.187×10^{-2}	3.345×10^{-2}	6.387×10^{-2}
	$\pm 5.65 \times 10^{-5}$	$\pm 8.98 \times 10^{-5}$	$\pm 8.77 \times 10^{-5}$	$\pm 1.20 \times 10^{-4}$	1.22×10^{-4}	$\pm 3.60 \times 10^{-4}$	$\pm 1.73 \times 10^{-3}$

Table 32: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the M/M/s model with s = 100 for actual delays in $(4\widehat{E[W\|W>0]}, 6\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	4.937×10^{-2}	8.655×10^{-2}	8.632×10^{-2}	8.943×10^{-2}	9.008×10^{-2}	1.088×10^{-1}	11.586
	$\pm 7.03 \times 10^{-3}$	$\pm 6.89 \times 10^{-3}$	$\pm 6.86 \times 10^{-3}$	$\pm 7.24 \times 10^{-3}$	$\pm 7.23 \times 10^{-3}$	$\pm 1.02 \times 10^{-2}$	± 1.25
0.98	2.479×10^{-2}	4.747×10^{-2}	4.725×10^{-2}	5.014×10^{-2}	5.057×10^{-2}	6.964×10^{-2}	3.542
	$\pm 1.75 \times 10^{-3}$	$\pm 3.06 \times 10^{-3}$	$\pm 3.04 \times 10^{-3}$	$\pm 3.09 \times 10^{-3}$	$\pm 3.13 \times 10^{-3}$	$\pm 3.67 \times 10^{-3}$	$\pm 4.31 \times 10^{-1}$
0.95	1.052×10^{-2}	2.037×10^{-2}	2.011×10^{-2}	2.350×10^{-2}	2.403×10^{-2}	5.039×10^{-2}	5.641×10^{-1}
	$\pm 2.30 \times 10^{-4}$	$\pm 6.30 \times 10^{-4}$	$\pm 6.16 \times 10^{-4}$	$\pm 8.16 \times 10^{-4}$	$\pm 8.00 \times 10^{-4}$	$\pm 3.34 \times 10^{-3}$	$\pm 2.70 \times 10^{-2}$
0.93	7.544×10^{-3}	1.515×10^{-2}	1.487×10^{-2}	1.870×10^{-2}	1.930×10^{-2}	5.204×10^{-2}	2.860×10^{-1}
	$\pm 1.97 \times 10^{-3}$	$\pm 3.08 \times 10^{-4}$	$\pm 2.90 \times 10^{-4}$	$\pm 4.28 \times 10^{-4}$	$\pm 4.52 \times 10^{-4}$	$\pm 3.18 \times 10^{-3}$	$\pm 8.01 \times 10^{-3}$
0.90	5.622×10^{-3}	1.105×10^{-2}	1.073×10^{-2}	1.531×10^{-2}	1.609×10^{-2}	5.089×10^{-2}	1.374×10^{-1}
	$\pm 2.07 \times 10^{-4}$	$\pm 3.82 \times 10^{-4}$	$\pm 3.75 \times 10^{-4}$	$\pm 6.07 \times 10^{-4}$	$\pm 6.62 \times 10^{-4}$	$\pm 2.52 \times 10^{-2}$	$\pm 6.74 \times 10^{-3}$

Table 33: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4\widehat{E[W|W>0]}, 6\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the M/M/s model with $s = 100$ *for actual delays* $> 6E[\widehat{W|W} > 0])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	3.361×10^{-2}	7.526×10^{-2}	7.521×10^{-2}	7.735×10^{-2}	7.776×10^{-2}	9.012×10^{-2}	7.902
	$\pm 1.21 \times 10^{-2}$	$\pm 2.61 \times 10^{-2}$	$\pm 2.60 \times 10^{-2}$	$\pm 2.63 \times 10^{-2}$	$\pm 2.64 \times 10^{-2}$	$\pm 2.30 \times 10^{-2}$	± 1.29
0.95	1.862×10^{-2}	3.353×10^{-2}	3.319×10^{-2}	3.755×10^{-2}	3.820×10^{-2}	6.823×10^{-2}	1.496
	$\pm 2.22 \times 10^{-3}$	$\pm 4.14 \times 10^{-3}$	$\pm 4.13 \times 10^{-3}$	$\pm 4.20 \times 10^{-3}$	$\pm 4.49 \times 10^{-3}$	$\pm 8.10 \times 10^{-3}$	$\pm 1.65 \times 10^{-1}$
0.93	1.306×10^{-2}	2.306×10^{-2}	2.274×10^{-2}	2.739×10^{-2}	2.803×10^{-2}	6.603×10^{-2}	7.362×10^{-1}
	$\pm 1.26 \times 10^{-3}$	$\pm 3.16 \times 10^{-3}$	$\pm 3.06 \times 10^{-3}$	$\pm 3.82 \times 10^{-3}$	$\pm 3.96 \times 10^{-3}$	$\pm 1.07 \times 10^{-2}$	$\pm 6.29 \times 10^{-2}$
0.90	1.042×10^{-2}	1.872×10^{-2}	1.828×10^{-2}	2.378×10^{-2}	2.438×10^{-2}	7.160×10^{-2}	0.3478
	$\pm 9.67 \times 10^{-4}$	$\pm 8.63 \times 10^{-4}$	$\pm 8.65 \times 10^{-4}$	$\pm 1.41 \times 10^{-3}$	$\pm 1.34 \times 10^{-3}$	$\pm 6.04 \times 10^{-3}$	$\pm 3.67 \times 10^{-2}$

Table 34: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) when delays are larger than $6E[\widehat{W|W} > 0]$. Each estimate is shown with the half width of the 95 percent confidence interval.

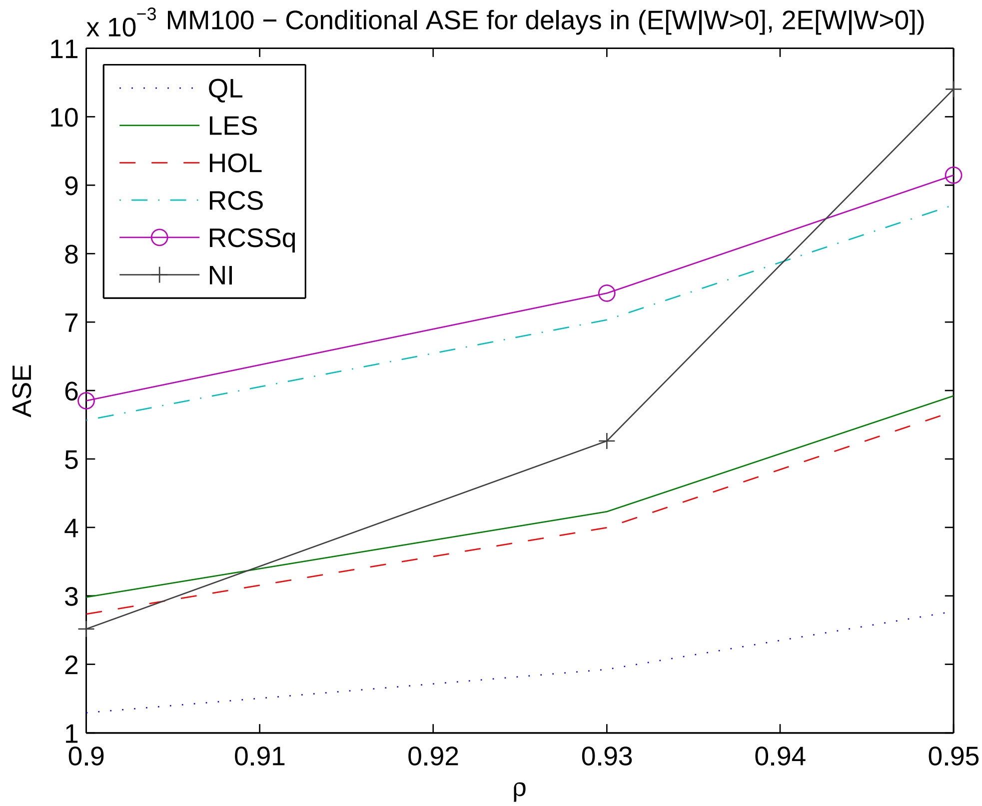

Figure 31: Conditional ASE for the alternative delay estimators in the $M/M/100$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

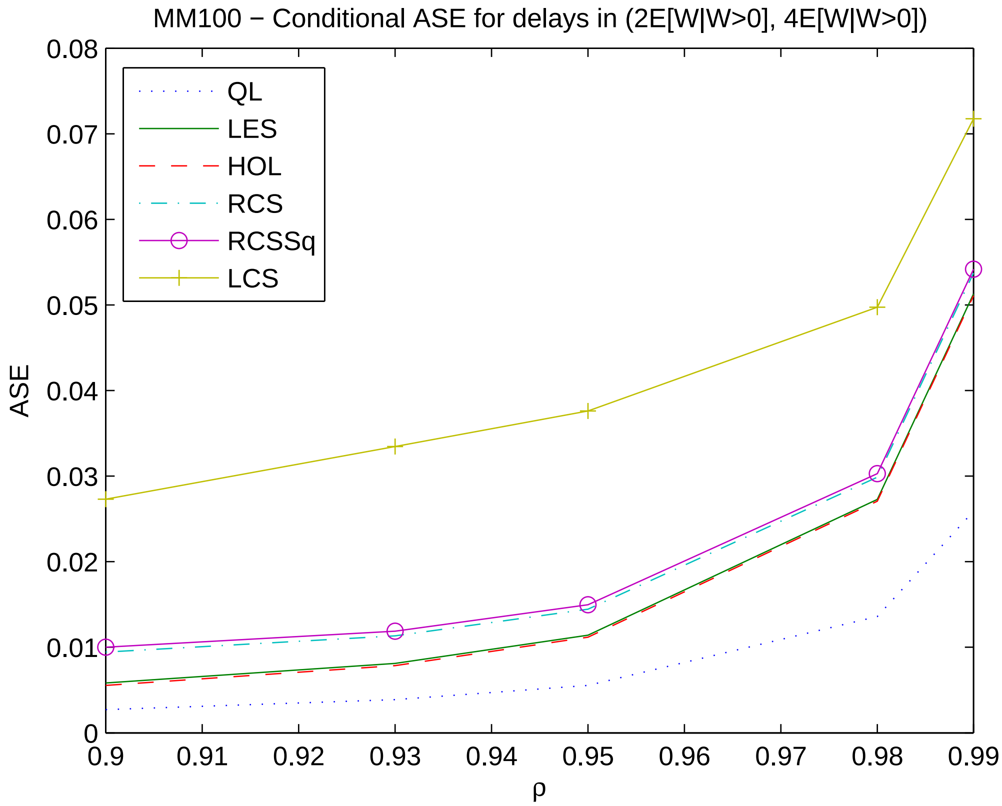

Figure 32: Conditional ASE for the alternative delay estimators in the $M/M/100$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

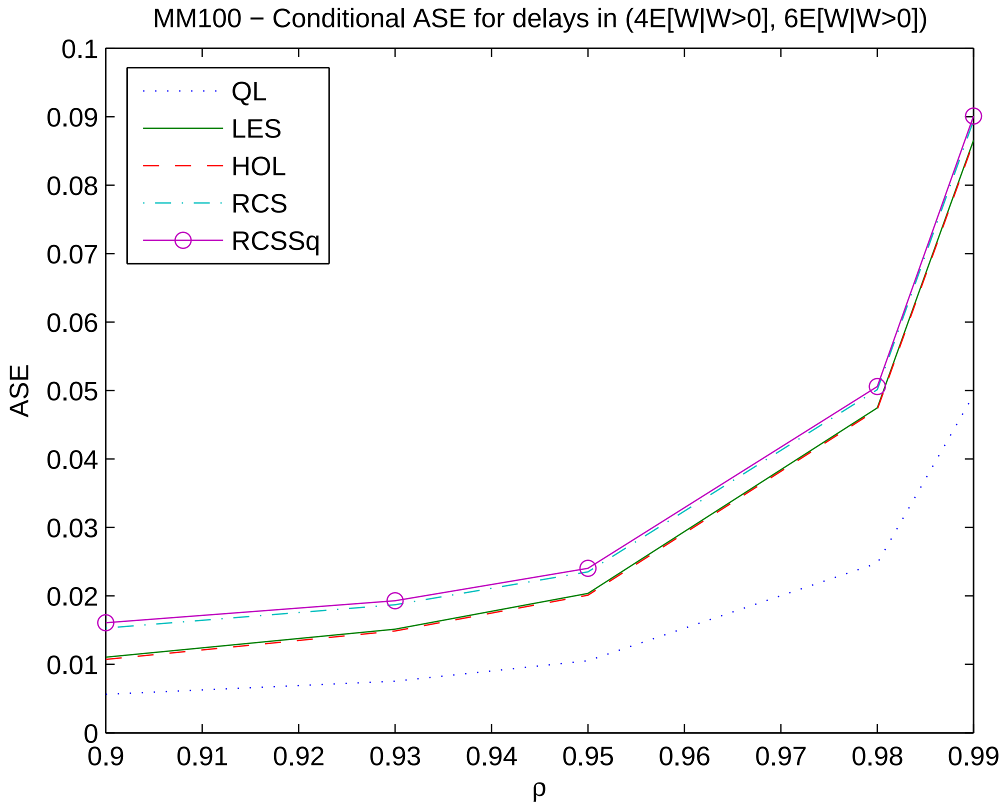

Figure 33: Conditional ASE for the alternative delay estimators in the $M/M/100$ model for actual delays in $(4E[\widehat{W|W}>0], 6E[\widehat{W|W}>0])$, as a function of the traffic intensity ρ

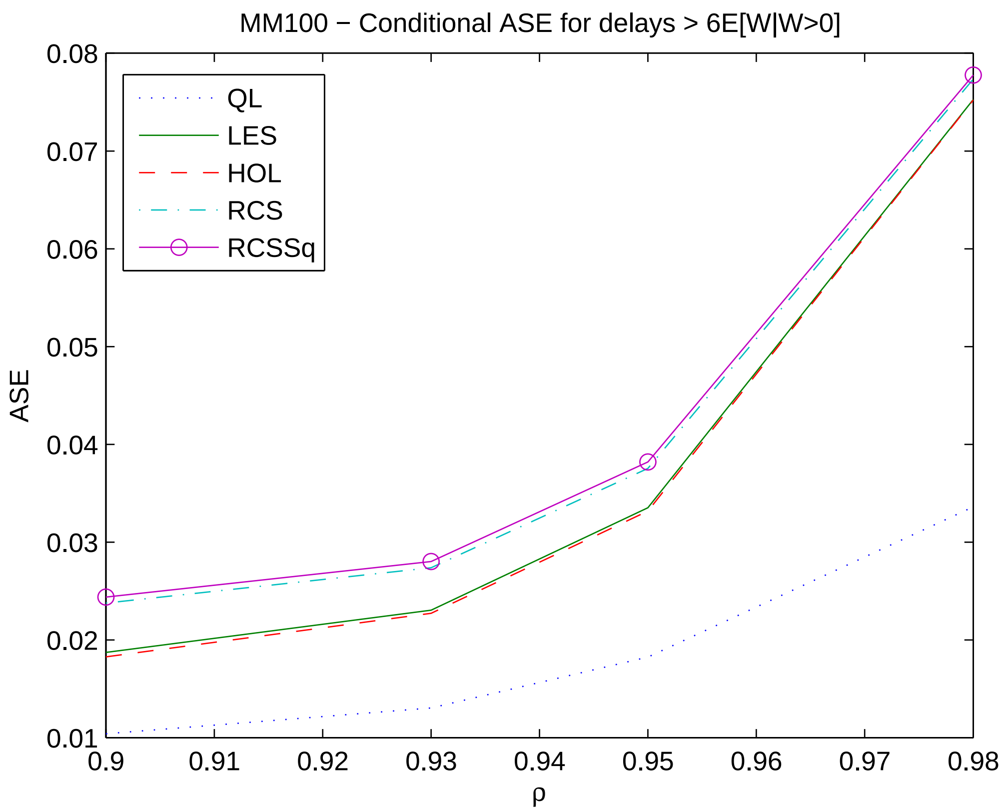

Figure 34: Conditional ASE for the alternative delay estimators in the $M/M/100$ model for actual delays larger than $6E[\widehat{W|W > 0}]$, as a function of the traffic intensity ρ

Conditional ASE in the $M/M/s$ model with $s = 400$ for actual delays in $(E[\widehat{W\|W > 0}], 2E[\widehat{W\|W > 0}])$								
ρ	$E[\widehat{W\|W > 0}]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	2.507×10^{-1}	8.808×10^{-4}	1.796×10^{-3}	1.782×10^{-3}	2.122×10^{-3}	2.177×10^{-3}	7.164×10^{-3}	1.607×10^{-2}
	$\pm 1.94 \times 10^{-2}$	$\pm 6.75 \times 10^{-5}$	$\pm 1.38 \times 10^{-4}$	$\pm 1.38 \times 10^{-4}$	$\pm 1.42 \times 10^{-4}$	$\pm 1.40 \times 10^{-4}$	$\pm 2.64 \times 10^{-4}$	$\pm 2.70 \times 10^{-3}$
0.98	1.267×10^{-1}	4.388×10^{-4}	9.035×10^{-4}	8.905×10^{-4}	1.228×10^{-4}	1.283×10^{-3}	5.733×10^{-3}	4.104×10^{-3}
	$\pm 4.967 \times 10^{-3}$	$\pm 1.77 \times 10^{-5}$	$\pm 3.73 \times 10^{-4}$	$\pm 3.72 \times 10^{-5}$	$\pm 3.70 \times 10^{-5}$	$\pm 3.70 \times 10^{-5}$	$\pm 1.87 \times 10^{-4}$	$\pm 3.04 \times 10^{-4}$
0.95	5.052×10^{-2}	1.710×10^{-4}	3.656×10^{-4}	3.518×10^{-4}	7.136×10^{-4}	7.628×10^{-4}	2.828×10^{-3}	6.487×10^{-4}
	$\pm 2.03 \times 10^{-3}$	$\pm 7.90 \times 10^{-6}$	$\pm 1.47 \times 10^{-5}$	$\pm 1.46 \times 10^{-5}$	$\pm 2.18 \times 10^{-5}$	$\pm 2.25 \times 10^{-5}$	$\pm 1.52 \times 10^{-4}$	$\pm 4.91 \times 10^{-5}$
0.93	3.680×10^{-2}	1.232×10^{-4}	2.713×10^{-4}	2.562×10^{-4}	6.078×10^{-4}	6.497×10^{-4}	1.937×10^{-3}	3.443×10^{-4}
	$\pm 1.50 \times 10^{-3}$	$\pm 5.77 \times 10^{-6}$	$\pm 1.16 \times 10^{-5}$	$\pm 1.15 \times 10^{-5}$	$\pm 2.07 \times 10^{-5}$	$\pm 2.25 \times 10^{-5}$	$\pm 1.27 \times 10^{-4}$	2.83×10^{-5}
0.90	2.531×10^{-2}	8.297×10^{-5}	1.920×10^{-4}	1.764×10^{-4}	4.872×10^{-4}	5.121×10^{-4}	1.117×10^{-3}	1.637×10^{-4}
	$\pm 1.31 \times 10^{-3}$	$\pm 4.83 \times 10^{-6}$	$\pm 9.48 \times 10^{-6}$	$\pm 9.43 \times 10^{-6}$	$\pm 1.92 \times 10^{-5}$	$\pm 2.27 \times 10^{-5}$	$\pm 9.57 \times 10^{-5}$	$\pm 1.51 \times 10^{-5}$

Table 35: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the M/M/s model with $s = 400$ *for actual delays in* $(2\widehat{E[W\|W>0]}, 4\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.736×10^{-3}	3.431×10^{-3}	3.417×10^{-3}	3.757×10^{-3}	3.814×10^{-3}	8.890×10^{-3}	2.079×10^{-1}
	$\pm 1.72 \times 10^{-4}$	$\pm 3.15 \times 10^{-4}$	$\pm 3.15 \times 10^{-4}$	$\pm 3.18 \times 10^{-4}$	$\pm 3.20 \times 10^{-4}$	$\pm 3.06 \times 10^{-4}$	4.05×10^{-2}
0.98	8.438×10^{-4}	1.721×10^{-3}	1.707×10^{-3}	2.054×10^{-3}	2.106×10^{-3}	7.654×10^{-3}	5.067×10^{-2}
	$\pm 4.25 \times 10^{-5}$	$\pm 8.22 \times 10^{-5}$	$\pm 8.25 \times 10^{-5}$	$\pm 8.04 \times 10^{-5}$	$\pm 8.26 \times 10^{-5}$	$\pm 2.74 \times 10^{-4}$	6.02×10^{-3}
0.95	3.460×10^{-4}	7.08×10^{-4}	6.928×10^{-4}	1.094×10^{-3}	1.162×10^{-3}	6.117×10^{-3}	7.913×10^{-3}
	$\pm 1.66 \times 10^{-5}$	$\pm 3.42 \times 10^{-5}$	$\pm 3.41 \times 10^{-5}$	$\pm 4.12 \times 10^{-5}$	$\pm 4.49 \times 10^{-5}$	2.43×10^{-4}	6.31×10^{-4}
0.93	2.475×10^{-4}	5.220×10^{-4}	5.060×10^{-4}	9.352×10^{-4}	1.003×10^{-3}	4.757×10^{-3}	4.283×10^{-3}
	$\pm 1.54 \times 10^{-5}$	$\pm 3.09 \times 10^{-5}$	$\pm 3.05 \times 10^{-5}$	$\pm 3.41 \times 10^{-5}$	$\pm 3.61 \times 10^{-5}$	2.07×10^{-4}	3.54×10^{-4}
0.90	1.753×10^{-4}	3.763×10^{-4}	3.587×10^{-4}	8.439×10^{-4}	9.150×10^{-4}	3.207×10^{-3}	2.01×10^{-3}
	$\pm 1.33 \times 10^{-5}$	$\pm 2.54 \times 10^{-5}$	$\pm 2.49 \times 10^{-5}$	$\pm 4.16 \times 10^{-5}$	$\pm 4.96 \times 10^{-5}$	$\pm 2.53 \times 10^{-4}$	$\pm 2.30 \times 10^{-4}$

Table 36: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W|W>0]}, 4\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the M/M/s model with s = 400 for actual delays in $(4\widehat{E[W\|W>0]}, 6\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	3.026×10^{-3}	5.939×10^{-3}	5.926×10^{-3}	6.296×10^{-3}	6.350×10^{-3}	1.194×10^{-2}	9.053×10^{-1}
	$\pm 4.09 \times 10^{-4}$	$\pm 6.95 \times 10^{-4}$	$\pm 6.96 \times 10^{-4}$	$\pm 7.16 \times 10^{-4}$	$\pm 7.15 \times 10^{-4}$	$\pm 1.41 \times 10^{-3}$	2.04×10^{-1}
0.98	1.465×10^{-3}	2.974×10^{-3}	2.961×10^{-3}	3.321×10^{-3}	3.366×10^{-3}	9.149×10^{-3}	2.156×10^{-1}
$\pm 1.91 \times 10^{-4}$	$\pm 2.53 \times 10^{-4}$	2.52×10^{-4}	$\pm 2.70 \times 10^{-4}$	2.75×10^{-4}	5.70×10^{-4}	1.97×10^{-2}	
0.95	6.489×10^{-4}	1.345×10^{-3}	1.325×10^{-3}	1.808×10^{-3}	1.892×10^{-3}	1.010×10^{-2}	3.573×10^{-2}
	$\pm 3.34 \times 10^{-5}$	$\pm 9.04 \times 10^{-5}$	9.07×10^{-5}	1.15×10^{-4}	1.11×10^{-4}	7.16×10^{-4}	3.59×10^{-3}
0.93	4.862×10^{-4}	9.684×10^{-4}	9.512×10^{-4}	1.394×10^{-3}	1.468×10^{-3}	9.316×10^{-3}	1.909×10^{-2}
	$\pm 3.87 \times 10^{-5}$	$\pm 7.86 \times 10^{-5}$	$\pm 7.45 \times 10^{-5}$	1.33×10^{-4}	1.28×10^{-4}	6.20×10^{-4}	1.52×10^{-3}
0.90	3.87×10^{-4}	7.00×10^{-4}	6.801×10^{-4}	1.282×10^{-3}	1.393×10^{-3}	7.896×10^{-3}	8.682×10^{-3}
$\pm 5.20 \times 10^{-5}$	$\pm 8.78 \times 10^{-5}$	$\pm 8.73 \times 10^{-5}$	$\pm 1.73 \times 10^{-4}$	$\pm 1.65 \times 10^{-4}$	$\pm 6.71 \times 10^{-4}$	$\pm 9.75 \times 10^{-4}$	

Table 37: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4\widehat{E[W|W>0]}, 6\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

ASE *in the* $M/M/s$ *model with* $s = 400$ *when delays* $> 6E[W\|W > 0]$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	9.815×10^{-4}	1.936×10^{-3}	1.916×10^{-3}	2.311×10^{-3}	2.310×10^{-3}	1.470×10^{-2}	8.352×10^{-2}
	$\pm 1.64 \times 10^{-4}$	$\pm 3.64 \times 10^{-4}$	$\pm 3.52 \times 10^{-4}$	$\pm 4.74 \times 10^{-4}$	$\pm 4.77 \times 10^{-4}$	$\pm 2.52 \times 10^{-3}$	$\pm 8.66 \times 10^{-3}$
0.93	7.0711×10^{-4}	1.359×10^{-3}	1.338×10^{-3}	1.886×10^{-3}	2.003×10^{-3}	1.450×10^{-2}	4.483×10^{-2}
	$\pm 1.61 \times 10^{-4}$	$\pm 3.10 \times 10^{-4}$	$\pm 3.03 \times 10^{-4}$	$\pm 5.09 \times 10^{-4}$	$\pm 5.33 \times 10^{-4}$	$\pm 3.65 \times 10^{-3}$	$\pm 4.43 \times 10^{-3}$
0.90	9.73×10^{-4}	1.78×10^{-3}	1.708×10^{-3}	2.359×10^{-3}	2.629×10^{-3}	1.457×10^{-2}	1.965×10^{-2}
	$\pm 2.84 \times 10^{-4}$	$\pm 5.16 \times 10^{-4}$	$\pm 4.79 \times 10^{-4}$	$\pm 6.85 \times 10^{-4}$	$\pm 7.33 \times 10^{-4}$	$\pm 2.27 \times 10^{-3}$	$\pm 3.17 \times 10^{-3}$

Table 38: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) when delays are larger than $6E[\widehat{W|W > 0}]$. Each estimate is shown with the half width of the 95 percent confidence interval.

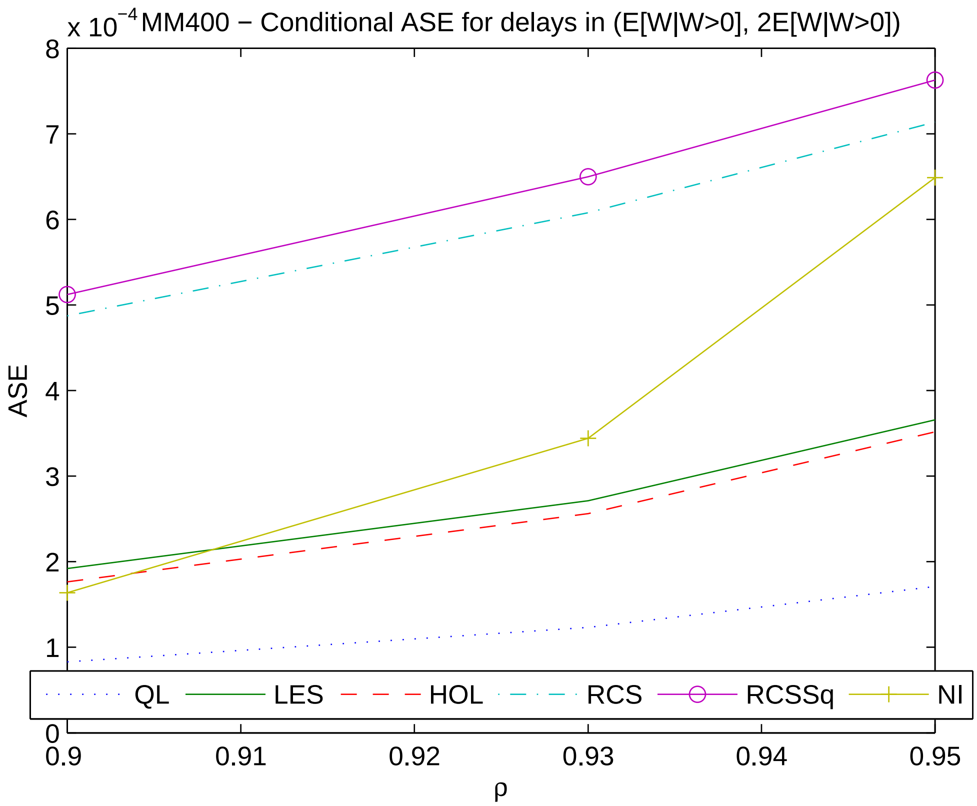

Figure 35: Conditional ASE for the alternative delay estimators in the $M/M/400$ model for actual delays in $(E[\widehat{W|W>0}], 2E[\widehat{W|W>0}])$, as a function of the traffic intensity ρ

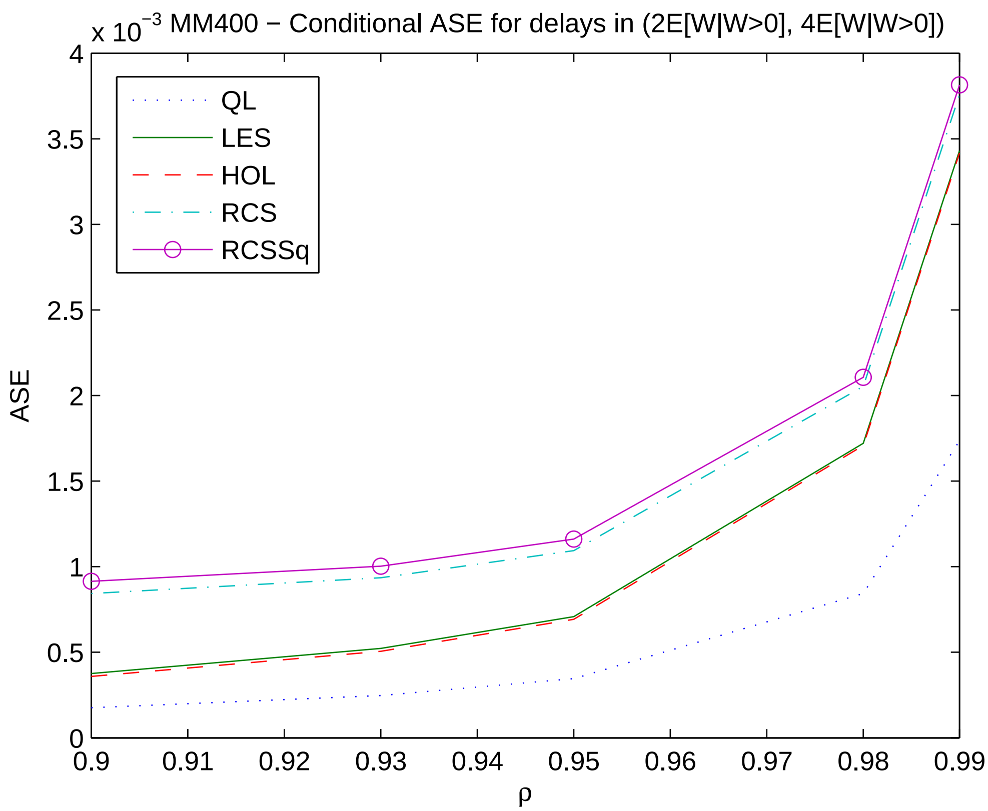

Figure 36: Conditional ASE for the alternative delay estimators in the $M/M/400$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

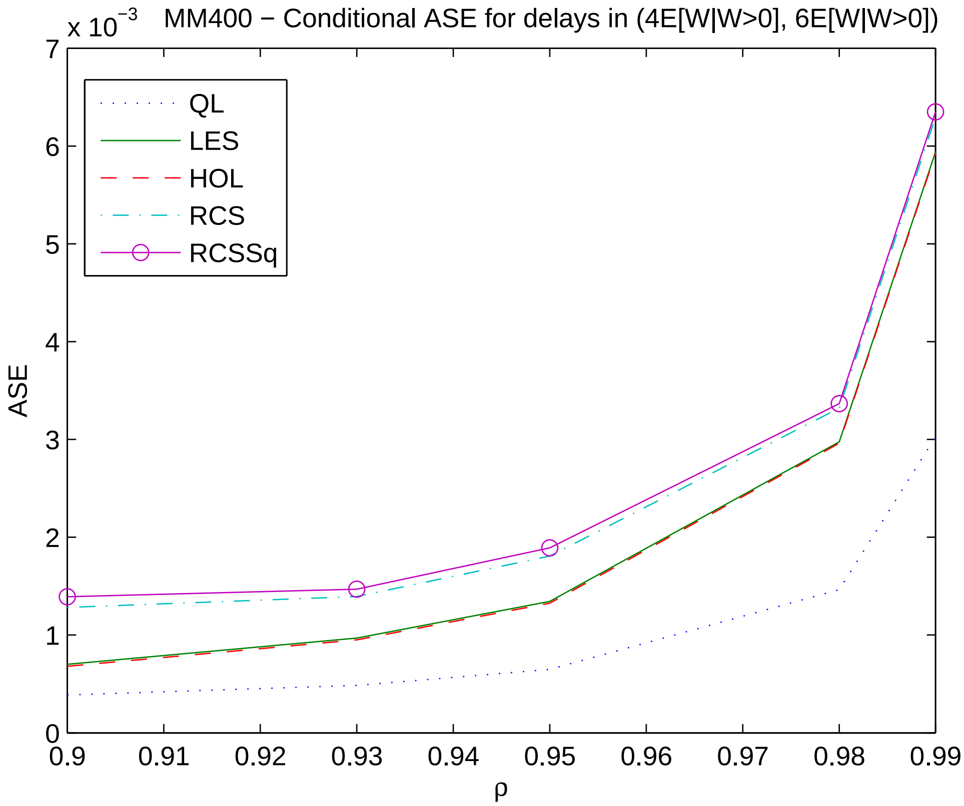

Figure 37: Conditional ASE for the alternative delay estimators in the $M/M/400$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

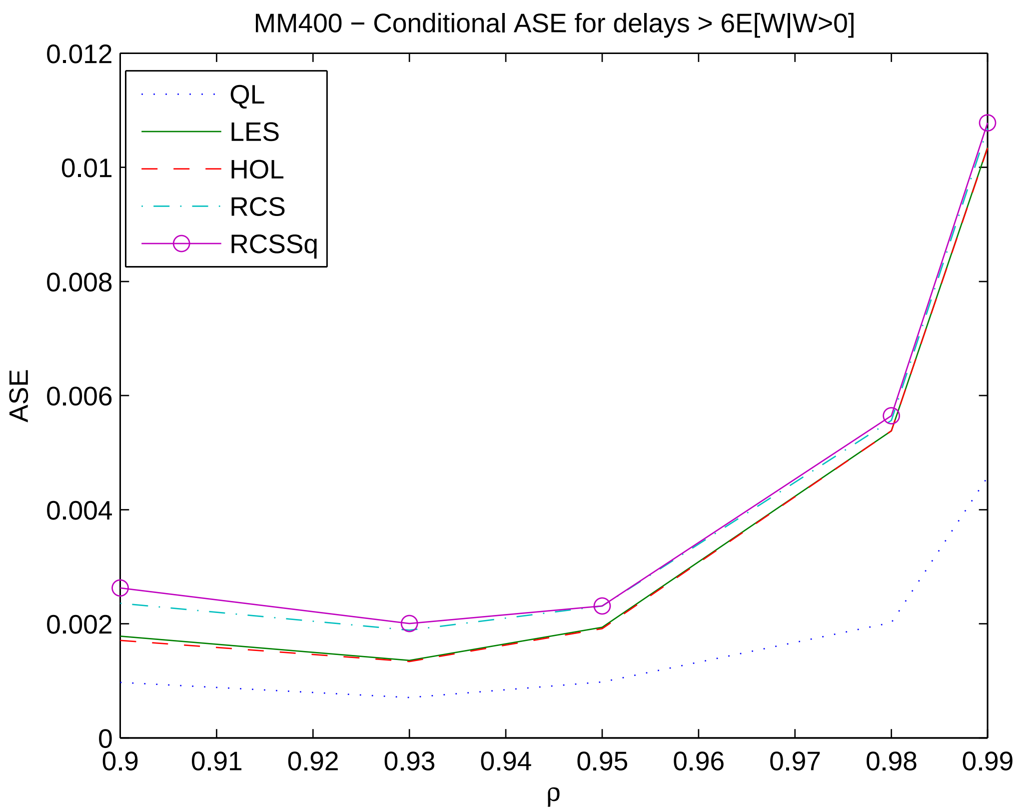

Figure 38: Conditional ASE for the alternative delay estimators in the $M/M/400$ model for actual delays larger than $6E[\widehat{W|W > 0}])$, as a function of the traffic intensity ρ

Conditional ASE in the M/M/s model with $s = 900$ *for actual delays in* $(E[\widehat{W\|W > 0}], 2E[\widehat{W\|W > 0}])$								
ρ	$E[\widehat{W\|W > 0}]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.193×10^{-1}	1.855×10^{-4}	3.774×10^{-4}	3.749×10^{-4}	4.721×10^{-4}	4.880×10^{-4}	2.677×10^{-3}	3.636×10^{-3}
	$\pm 8.27 \times 10^{-3}$	$\pm 1.40 \times 10^{-5}$	$\pm 2.67 \times 10^{-5}$	$\pm 2.67 \times 10^{-5}$	$\pm 2.78 \times 10^{-5}$	$\pm 2.80 \times 10^{-5}$	$\pm 8.31 \times 10^{-5}$	$\pm 6.01 \times 10^{-4}$
0.98	5.690×10^{-2}	8.814×10^{-5}	1.812×10^{-4}	1.787×10^{-4}	2.788×10^{-4}	2.954×10^{-4}	1.943×10^{-3}	8.292×10^{-4}
	$\pm 1.86 \times 10^{-3}$	$\pm 2.69 \times 10^{-6}$	$\pm 4.79 \times 10^{-6}$	$\pm 4.79 \times 10^{-6}$	$\pm 4.42 \times 10^{-6}$	$\pm 4.58 \times 10^{-6}$	$\pm 5.16 \times 10^{-5}$	$\pm 5.17 \times 10^{-5}$
0.95	2.235×10^{-2}	3.357×10^{-5}	7.204×10^{-5}	6.924×10^{-5}	1.744×10^{-4}	1.882×10^{-4}	7.396×10^{-4}	1.264×10^{-4}
	$\pm 1.52 \times 10^{-3}$	$\pm 2.47 \times 10^{-6}$	$\pm 4.99 \times 10^{-6}$	$\pm 4.98 \times 10^{-6}$	$\pm 7.47 \times 10^{-6}$	$\pm 8.49 \times 10^{-6}$	$\pm 7.81 \times 10^{-5}$	$\pm 1.55 \times 10^{-5}$
0.93	1.573×10^{-2}	2.347×10^{-5}	5.26×10^{-5}	4.97×10^{-5}	1.460×10^{-4}	1.554×10^{-4}	4.367×10^{-4}	6.386×10^{-5}
	$\pm 1.66 \times 10^{-3}$	$\pm 2.80 \times 10^{-6}$	$\pm 5.61 \times 10^{-6}$	$\pm 5.52 \times 10^{-6}$	$\pm 1.20 \times 10^{-5}$	$\pm 1.40 \times 10^{-5}$	$\pm 7.68 \times 10^{-5}$	$\pm 1.28 \times 10^{-5}$

Table 39: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $M/M/s$ model with $s = 900$ for actual delays in $(2\widehat{E[W \mid W > 0]}, 4\widehat{E[W \mid W > 0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	3.533×10^{-4}	7.072×10^{-4}	7.046×10^{-4}	7.981×10^{-4}	8.137×10^{-4}	3.079×10^{-3}	4.602×10^{-2}
	$\pm 2.85 \times 10^{-5}$	$\pm 6.37 \times 10^{-5}$	$\pm 6.37 \times 10^{-5}$	$\pm 6.48 \times 10^{-5}$	$\pm 6.55 \times 10^{-5}$	$\pm 1.83 \times 10^{-4}$	$\pm 8.93 \times 10^{-3}$
0.98	1.690×10^{-4}	3.438×10^{-4}	3.411×10^{-4}	4.432×10^{-4}	4.604×10^{-4}	3.048×10^{-3}	1.005×10^{-2}
	$\pm 8.53 \times 10^{-6}$	$\pm 1.26 \times 10^{-5}$	$\pm 1.26 \times 10^{-5}$	$\pm 1.35 \times 10^{-5}$	$\pm 1.36 \times 10^{-5}$	$\pm 9.60 \times 10^{-5}$	$\pm 7.72 \times 10^{-4}$
0.95	6.793×10^{-5}	1.403×10^{-4}	1.374×10^{-4}	2.563×10^{-4}	2.779×10^{-4}	1.905×10^{-3}	1.602×10^{-3}
	$\pm 4.61 \times 10^{-6}$	$\pm 1.12 \times 10^{-5}$	$\pm 1.11 \times 10^{-5}$	$\pm 1.02 \times 10^{-5}$	$\pm 1.08 \times 10^{-5}$	$\pm 1.89 \times 10^{-4}$	$\pm 2.36 \times 10^{-4}$
0.93	4.85×10^{-5}	1.054×10^{-4}	1.018×10^{-4}	2.390×10^{-4}	2.608×10^{-4}	1.300×10^{-3}	8.177×10^{-4}
	$\pm 5.45 \times 10^{-6}$	$\pm 1.25 \times 10^{-5}$	$\pm 1.23 \times 10^{-5}$	$\pm 2.02 \times 10^{-5}$	$\pm 2.09 \times 10^{-5}$	$\pm 2.27 \times 10^{-4}$	$\pm 1.75 \times 10^{-4}$

Table 40: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W \mid W > 0]}, 4\widehat{E[W \mid W > 0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $M/M/s$ model with $s = 900$ for actual delays in $(4\widehat{E[W\|W>0]}, 6\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	6.086×10^{-4}	1.287×10^{-3}	1.284×10^{-3}	1.394×10^{-3}	1.410×10^{-3}	4.097×10^{-3}	1.949×10^{-1}
	$\pm 6.72 \times 10^{-5}$	$\pm 1.15 \times 10^{-4}$	$\pm 1.16 \times 10^{-4}$	$\pm 1.16 \times 10^{-4}$	$\pm 1.16 \times 10^{-4}$	$\pm 3.10 \times 10^{-4}$	3.86×10^{-2}
0.98	3.200×10^{-4}	6.322×10^{-4}	6.296×10^{-4}	7.302×10^{-4}	7.472×10^{-4}	3.879×10^{-3}	4.528×10^{-2}
	1.845×10^{-5}	4.808×10^{-5}	4.776×10^{-5}	5.709×10^{-5}	5.726×10^{-5}	3.061×10^{-4}	3.584×10^{-3}
0.95	1.211×10^{-4}	2.370×10^{-4}	2.342×10^{-4}	3.600×10^{-4}	3.812×10^{-4}	3.829×10^{-3}	6.899×10^{-3}
	$\pm 1.48 \times 10^{-5}$	$\pm 3.44 \times 10^{-5}$	$\pm 3.38 \times 10^{-5}$	$\pm 5.29 \times 10^{-5}$	$\pm 5.62 \times 10^{-5}$	$\pm 3.50 \times 10^{-4}$	$\pm 1.01 \times 10^{-3}$
0.93	9.350×10^{-5}	1.854×10^{-4}	1.812×10^{-4}	3.536×10^{-4}	3.815×10^{-4}	3.078×10^{-3}	3.365×10^{-3}
	$\pm 2.01 \times 10^{-5}$	$\pm 2.15 \times 10^{-5}$	$\pm 2.12 \times 10^{-5}$	$\pm 4.61 \times 10^{-5}$	$\pm 4.82 \times 10^{-5}$	$\pm 4.76 \times 10^{-4}$	$\pm 6.01 \times 10^{-4}$

Table 41: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4\widehat{E[W|W>0]}, 6\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

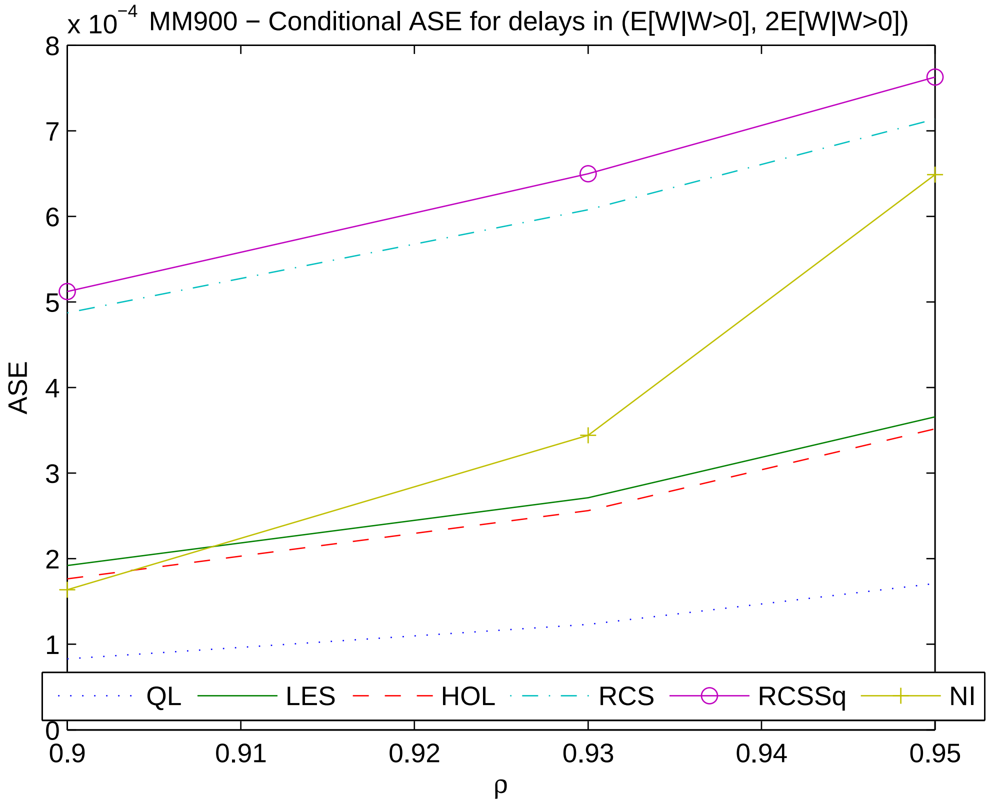

Figure 39: Conditional ASE for the alternative delay estimators in the $M/M/900$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

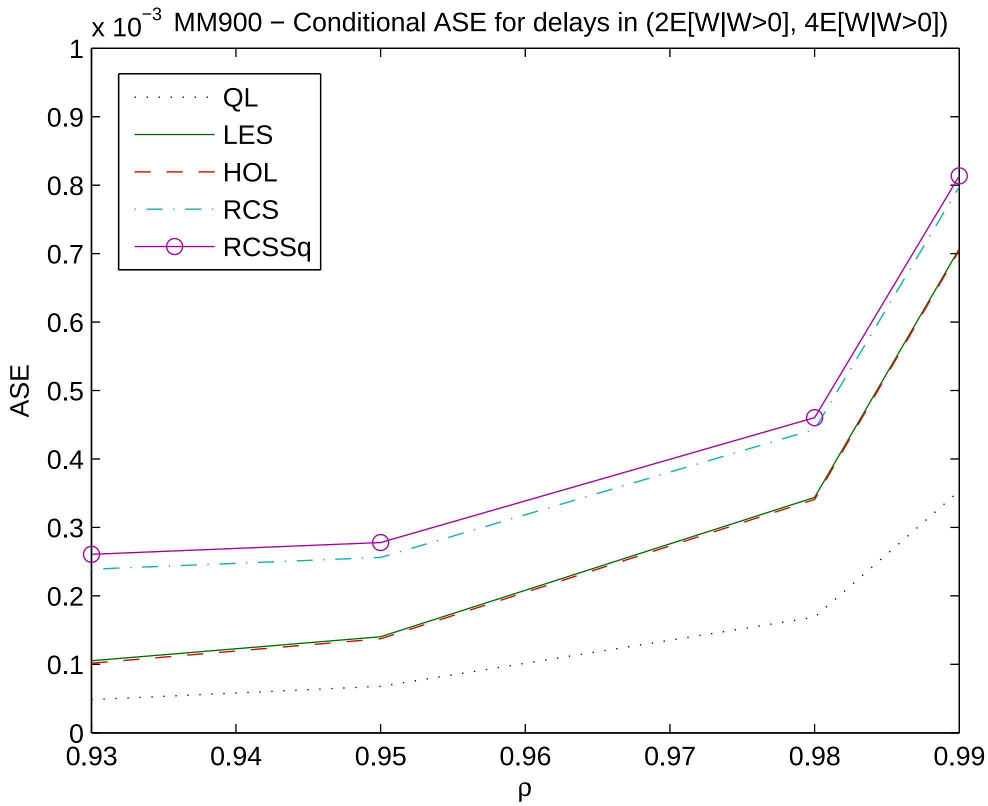

Figure 40: Conditional ASE for the alternative delay estimators in the $M/M/900$ model for actual delays in $(2E[\widehat{W|W > 0}], 4E[\widehat{W|W > 0}])$, as a function of the traffic intensity ρ

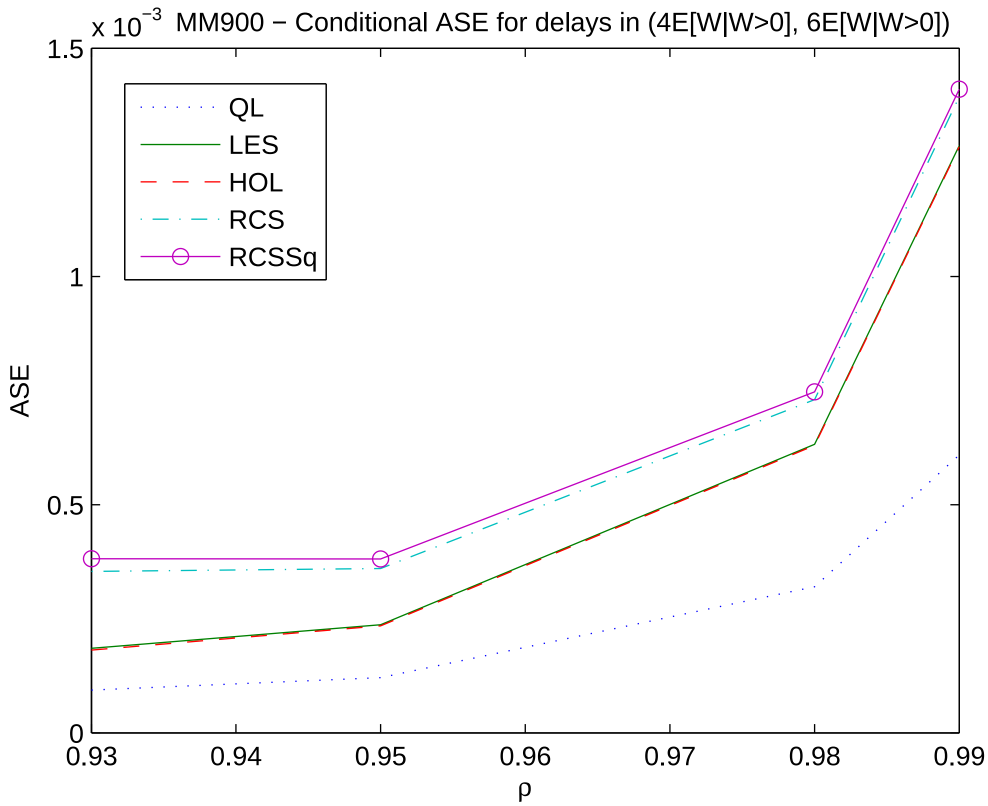

Figure 41: Conditional ASE for the alternative delay estimators in the $M/M/400$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

Conditional ASE in the D/M/s model with $s = 100$ *for actual delays in* $(E[\widehat{W \mid W} > 0], 2E[\widehat{W \mid W} > 0])$								
ρ	$E[\widehat{W \mid W} > 0]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	4.812×10^{-1}	6.649×10^{-3}	6.918×10^{-3}	6.814×10^{-3}	8.209×10^{-3}	8.415×10^{-3}	1.759×10^{-2}	5.764×10^{-2}
	$\pm 1.92 \times 10^{-2}$	$\pm 2.84 \times 10^{-4}$	$\pm 2.91 \times 10^{-4}$	$\pm 2.89 \times 10^{-4}$	$\pm 3.09 \times 10^{-4}$	$\pm 3.13 \times 10^{-4}$	$\pm 4.15 \times 10^{-4}$	$\pm 3.90 \times 10^{-3}$
0.98	2.475×10^{-1}	3.376×10^{-3}	3.655×10^{-3}	3.544×10^{-3}	5.026×10^{-3}	5.246×10^{-3}	1.439×10^{-2}	1.547×10^{-2}
	$\pm 8.30 \times 10^{-3}$	$\pm 1.17 \times 10^{-4}$	$\pm 1.22 \times 10^{-4}$	$\pm 1.21 \times 10^{-4}$	$\pm 1.28 \times 10^{-4}$	$\pm 1.32 \times 10^{-4}$	$\pm 3.26 \times 10^{-4}$	$\pm 9.72 \times 10^{-4}$
0.95	1.003×10^{-1}	1.307×10^{-3}	1.605×10^{-3}	1.476×10^{-3}	3.111×10^{-3}	3.314×10^{-3}	8.661×10^{-3}	2.567×10^{-3}
	$\pm 2.04 \times 10^{-3}$	$\pm 2.71 \times 10^{-5}$	$\pm 3.09 \times 10^{-5}$	$\pm 2.95 \times 10^{-5}$	$\pm 5.62 \times 10^{-5}$	$\pm 6.14 \times 10^{-5}$	$\pm 2.24 \times 10^{-4}$	$\pm 9.04 \times 10^{-5}$
0.93	7.242×10^{-2}	9.116×10^{-4}	1.223×10^{-3}	1.080×10^{-3}	2.650×10^{-3}	2.810×10^{-3}	6.259×10^{-3}	1.332×10^{-3}
	$\pm 1.29 \times 10^{-3}$	$\pm 1.59 \times 10^{-5}$	$\pm 1.77 \times 10^{-5}$	$\pm 1.79 \times 10^{-5}$	$\pm 3.96 \times 10^{-5}$	$\pm 4.07 \times 10^{-5}$	$\pm 1.45 \times 10^{-4}$	$\pm 4.11 \times 10^{-5}$
0.90	5.173×10^{-2}	6.199×10^{-4}	9.405×10^{-4}	7.757×10^{-4}	2.152×10^{-3}	2.254×10^{-3}	4.092×10^{-3}	6.811×10^{-4}
	$\pm 1.25 \times 10^{-3}$	$\pm 1.74 \times 10^{-5}$	2.29×10^{-5}	$\pm 2.16 \times 10^{-5}$	$\pm 5.99 \times 10^{-5}$	$\pm 6.53 \times 10^{-5}$	$\pm 1.60 \times 10^{-4}$	$\pm 3.08 \times 10^{-5}$

Table 42: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W \mid W} > 0], 2E[\widehat{W \mid W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the D/M/s model with s = 100 for actual delays in $(2\widehat{E[W\|W>0]}, 4\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.304×10^{-2}	1.332×10^{-2}	1.321×10^{-2}	1.461×10^{-2}	1.482×10^{-2}	2.436×10^{-2}	7.567×10^{-1}
	$\pm 9.16 \times 10^{-4}$	$\pm 9.33 \times 10^{-4}$	9.27×10^{-4}	9.95×10^{-4}	1.001×10^{-3}	1.503×10^{-3}	8.45×10^{-2}
0.98	6.635×10^{-3}	6.901×10^{-3}	6.786×10^{-3}	8.289×10^{-3}	8.516×10^{-3}	1.919×10^{-2}	1.90×10^{-1}
	$\pm 2.89 \times 10^{-4}$	$\pm 2.88 \times 10^{-4}$	$\pm 2.86 \times 10^{-4}$	$\pm 3.29 \times 10^{-4}$	$\pm 3.31 \times 10^{-4}$	$\pm 6.63 \times 10^{-4}$	$\pm 1.30 \times 10^{-2}$
0.95	2.729×10^{-3}	3.026×10^{-3}	2.887×10^{-3}	4.782×10^{-3}	5.065×10^{-3}	1.637×10^{-2}	3.160×10^{-2}
	$\pm 6.39 \times 10^{-5}$	$\pm 6.55 \times 10^{-5}$	$\pm 6.336 \times 10^{-5}$	$\pm 7.82 \times 10^{-5}$	$\pm 8.24 \times 10^{-5}$	$\pm 4.02 \times 10^{-4}$	$\pm 1.48 \times 10^{-3}$
0.93	1.988×10^{-3}	2.307×10^{-3}	2.149×10^{-3}	4.281×10^{-3}	4.584×10^{-3}	1.403×10^{-2}	1.647×10^{-2}
	$\pm 4.51 \times 10^{-5}$	$\pm 4.56 \times 10^{-5}$	$\pm 4.17 \times 10^{-5}$	$\pm 8.43 \times 10^{-5}$	$\pm 9.14 \times 10^{-5}$	$\pm 3.66 \times 10^{-4}$	$\pm 5.47 \times 10^{-4}$
0.9	1.436×10^{-3}	1.796×10^{-3}	1.603×10^{-3}	3.998×10^{-3}	4.264×10^{-3}	1.084×10^{-2}	8.367×10^{-3}
	$\pm 4.63 \times 10^{-5}$	$\pm 4.57 \times 10^{-5}$	$\pm 4.03 \times 10^{-5}$	1.032×10^{-4}	1.02×10^{-4}	$\pm 4.49 \times 10^{-4}$	$\pm 3.90 \times 10^{-4}$

Table 43: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W|W>0]}, 4\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the D/M/s model with s = 100 for actual delays in $(4\widehat{E[W\|W>0]}, 6\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	2.335×10^{-2}	2.309×10^{-2}	2.299×10^{-2}	2.442×10^{-2}	2.462×10^{-2}	3.432×10^{-2}	3.182
	$\pm 1.15 \times 10^{-3}$	$\pm 1.20 \times 10^{-3}$	$\pm 1.09 \times 10^{-3}$	$\pm 1.15 \times 10^{-3}$	$\pm 1.15 \times 10^{-3}$	$\pm 1.63 \times 10^{-3}$	$\pm 2.28 \times 10^{-1}$
0.98	1.277×10^{-2}	1.264×10^{-2}	1.253×10^{-2}	1.408×10^{-2}	1.432×10^{-2}	2.511×10^{-2}	8.800×10^{-1}
	$\pm 5.86 \times 10^{-4}$	$\pm 5.72 \times 10^{-4}$	$\pm 5.69 \times 10^{-4}$	$\pm 6.324 \times 10^{-4}$	$\pm 6.23 \times 10^{-4}$	$\pm 1.103 \times 10^{-3}$	$\pm 8.20 \times 10^{-2}$
0.95	5.707×10^{-3}	5.547×10^{-3}	5.390×10^{-3}	7.528×10^{-3}	7.821×10^{-3}	2.517×10^{-2}	1.384×10^{-1}
	$\pm 3.50 \times 10^{-4}$	$\pm 3.16 \times 10^{-4}$	$\pm 3.06 \times 10^{-4}$	$\pm 5.68 \times 10^{-4}$	$\pm 5.78 \times 10^{-4}$	$\pm 1.49 \times 10^{-3}$	$\pm 5.05 \times 10^{-3}$
0.93	4.426×10^{-3}	4.287×10^{-3}	4.107×10^{-3}	6.576×10^{-3}	6.980×10^{-3}	2.630×10^{-2}	7.239×10^{-2}
	$\pm 1.78 \times 10^{-4}$	$\pm 1.63 \times 10^{-4}$	$\pm 1.51 \times 10^{-4}$	$\pm 4.14 \times 10^{-4}$	$\pm 4.21 \times 10^{-4}$	$\pm 1.60 \times 10^{-3}$	$\pm 2.68 \times 10^{-3}$
0.90	3.394×10^{-3}	3.3197×10^{-3}	3.092×10^{-3}	6.214×10^{-3}	6.653×10^{-3}	2.363×10^{-2}	3.709×10^{-2}
	$\pm 1.50 \times 10^{-4}$	$\pm 1.49 \times 10^{-4}$	$\pm 1.38 \times 10^{-4}$	$\pm 3.69 \times 10^{-4}$	$\pm 4.02 \times 10^{-4}$	$\pm 1.34 \times 10^{-3}$	$\pm 1.65 \times 10^{-3}$

Table 44: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4\widehat{E[W|W>0]}, 6\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $D/M/s$ model with $s = 100$ for actual delays $> 6\widehat{E[W\|W > 0]}$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.965×10^{-2}	1.791×10^{-2}	1.780×10^{-2}	1.952×10^{-2}	1.981×10^{-2}	3.167×10^{-2}	2.0537
	$\pm 3.03 \times 10^{-3}$	$\pm 2.87 \times 10^{-3}$	$\pm 2.87 \times 10^{-3}$	$\pm 3.41 \times 10^{-3}$	$\pm 3.42 \times 10^{-3}$	6.83×10^{-3}	1.50×10^{-1}
0.95	9.421×10^{-3}	7.954×10^{-3}	7.804×10^{-3}	9.960×10^{-3}	1.032×10^{-2}	2.970×10^{-2}	3.550×10^{-1}
	$\pm 4.31 \times 10^{-4}$	$\pm 4.46 \times 10^{-4}$	$\pm 4.12 \times 10^{-4}$	$\pm 1.05 \times 10^{-3}$	$\pm 1.16 \times 10^{-3}$	$\pm 6.70 \times 10^{-3}$	$\pm 2.26 \times 10^{-2}$
0.93	8.018×10^{-3}	6.65×10^{-3}	6.439×10^{-3}	9.382×10^{-3}	9.711×10^{-3}	3.527×10^{-2}	1.849×10^{-2}
	$\pm 5.23 \times 10^{-4}$	$\pm 5.67 \times 10^{-4}$	$\pm 5.21 \times 10^{-4}$	$\pm 1.00 \times 10^{-3}$	$\pm 1.05 \times 10^{-3}$	$\pm 6.92 \times 10^{-3}$	$\pm 1.84 \times 10^{-2}$
0.90	6.789×10^{-3}	5.353×10^{-3}	5.103×10^{-3}	9.406×10^{-3}	9.994×10^{-3}	3.986×10^{-2}	9.402×10^{-2}
	$\pm 5.85 \times 10^{-4}$	$\pm 6.46 \times 10^{-4}$	$\pm 5.79 \times 10^{-4}$	$\pm 1.40 \times 10^{-3}$	$\pm 1.53 \times 10^{-3}$	$\pm 5.39 \times 10^{-3}$	$\pm 1.52 \times 10^{-2}$

Table 45: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) for delays larger than $6\widehat{E[W|W > 0]}$). Each estimate is shown with the half width of the 95 percent confidence interval.

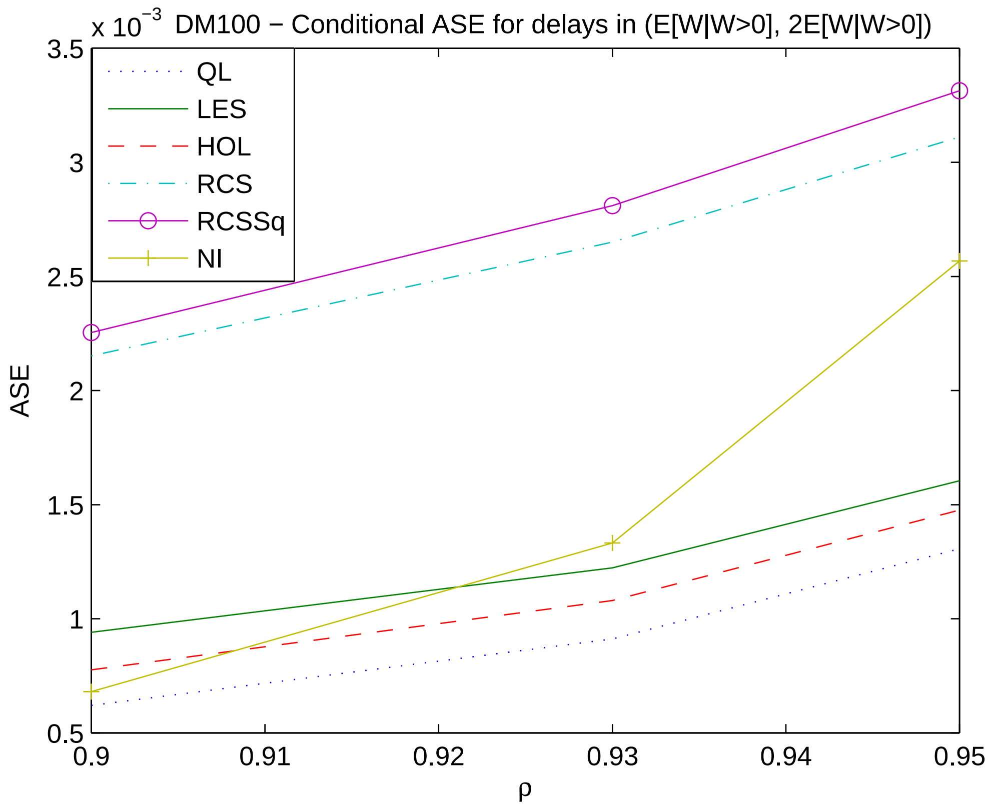

Figure 42: Conditional ASE for the alternative delay estimators in the $D/M/100$ model for actual delays in $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$, as a function of the traffic intensity ρ

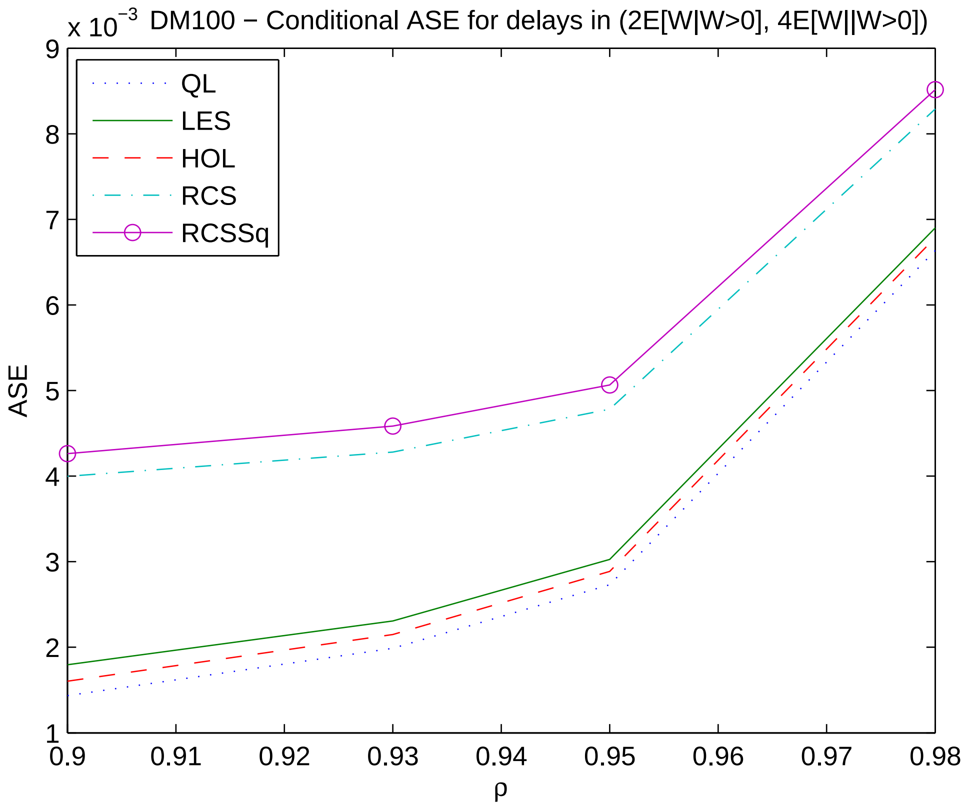

Figure 43: Conditional ASE for the alternative delay estimators in the $D/M/100$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

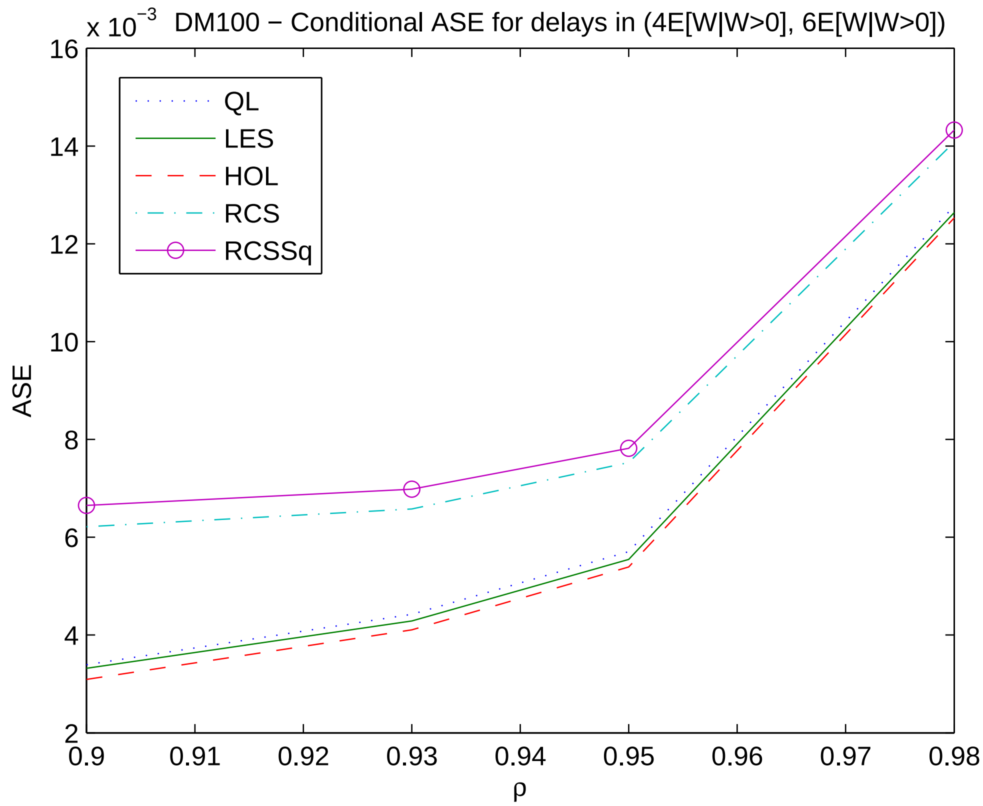

Figure 44: Conditional ASE for the alternative delay estimators in the $D/M/100$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

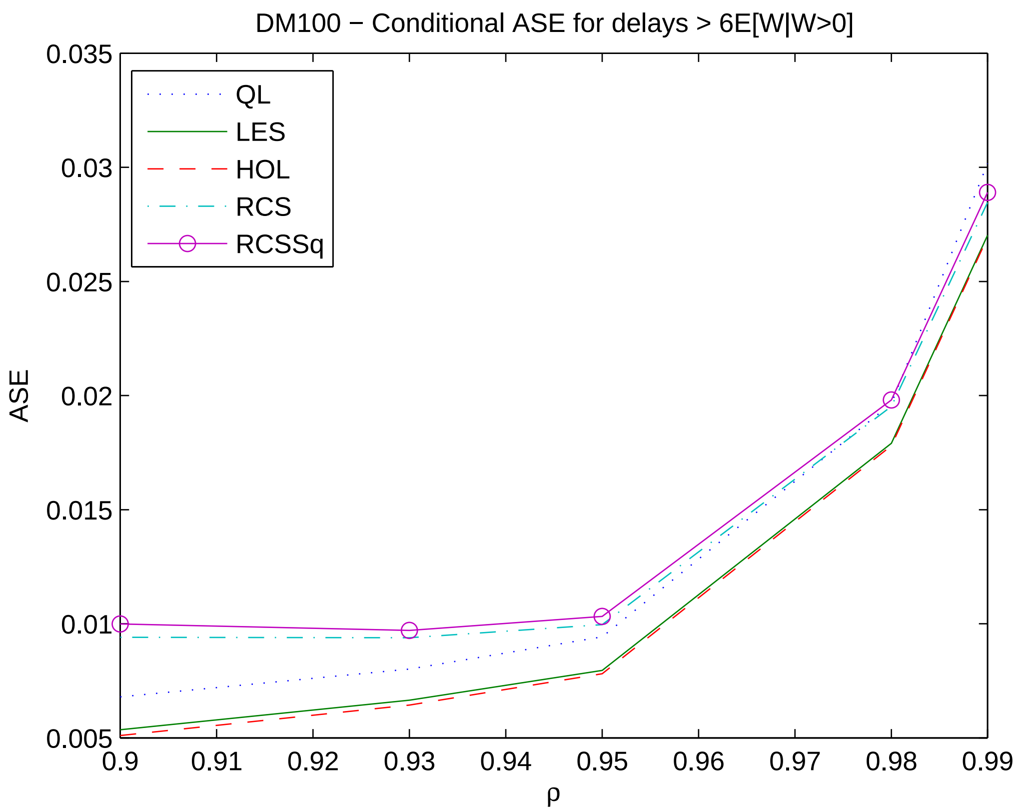

Figure 45: Conditional ASE for the alternative delay estimators in the $D/M/100$ model for actual delays larger than $6\widehat{E[W|W > 0]}$, as a function of the traffic intensity ρ

Conditional ASE in the $D/M/s$ model with $s = 400$ for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$

ρ	$E[\widehat{W\|W} > 0]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.212×10^{-1}	4.197×10^{-4}	4.367×10^{-4}	4.301×10^{-4}	5.988×10^{-4}	6.263×10^{-4}	3.040×10^{-3}	
	$\pm 5.38 \times 10^{-3}$	$\pm 1.82 \times 10^{-5}$	$\pm 1.85 \times 10^{-5}$	$\pm 1.85 \times 10^{-5}$	$\pm 2.01 \times 10^{-5}$	$\pm 2.02 \times 10^{-5}$	$\pm 8.88 \times 10^{-5}$	
0.98	6.214×10^{-2}	2.130×10^{-4}	2.304×10^{-4}	2.235×10^{-4}	4.049×10^{-4}	4.339×10^{-4}	2.286×10^{-3}	
	$\pm 2.56 \times 10^{-3}$	$\pm 9.31 \times 10^{-6}$	$\pm 9.61 \times 10^{-6}$	$\pm 9.54 \times 10^{-6}$	$\pm 1.21 \times 10^{-5}$	$\pm 1.22 \times 10^{-5}$	$\pm 1.00 \times 10^{-4}$	
0.95	2.497×10^{-2}	8.154×10^{-5}	1.003×10^{-4}	9.215×10^{-5}	2.778×10^{-4}	2.982×10^{-4}	9.087×10^{-4}	
	$\pm 1.18 \times 10^{-3}$	$\pm 4.07 \times 10^{-6}$	$\pm 4.29 \times 10^{-6}$	$\pm 4.35 \times 10^{-6}$	$\pm 8.98 \times 10^{-6}$	$\pm 9.30 \times 10^{-6}$	$\pm 6.55 \times 10^{-5}$	
0.93	1.834×10^{-2}	5.830×10^{-5}	7.804×10^{-5}	6.897×10^{-5}	2.367×10^{-4}	2.501×10^{-4}	5.798×10^{-4}	
	$\pm 8.68 \times 10^{-4}$	$\pm 3.25 \times 10^{-6}$	$\pm 3.57 \times 10^{-6}$	$\pm 3.60 \times 10^{-6}$	$\pm 1.08 \times 10^{-5}$	$\pm 1.16 \times 10^{-5}$	$\pm 4.57 \times 10^{-5}$	
0.9	1.258×10^{-2}	3.700×10^{-5}	5.615×10^{-5}	4.621×10^{-5}	1.738×10^{-4}	1.818×10^{-4}	3.035×10^{-4}	
	$\pm 1.478 \times 10^{-3}$	$\pm 5.029 \times 10^{-6}$	$\pm 6.28 \times 10^{-6}$	$\pm 6.21 \times 10^{-6}$	$\pm 2.41 \times 10^{-5}$	$\pm 2.80 \times 10^{-5}$	$\pm 6.97 \times 10^{-5}$	

Table 46: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the D/M/s model with $s = 400$ *for actual delays in* $(2E[\widehat{W \vert W > 0}], 4E[\widehat{W \vert W > 0}])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	8.373×10^{-4}	8.547×10^{-4}	8.479×10^{-4}	1.021×10^{-3}	1.050×10^{-3}	3.770×10^{-3}	4.858×10^{-2}
	$\pm 5.22 \times 10^{-5}$	$\pm 5.24 \times 10^{-5}$	$\pm 5.23 \times 10^{-5}$	$\pm 5.41 \times 10^{-5}$	$\pm 5.59 \times 10^{-5}$	$\pm 1.85 \times 10^{-4}$	$\pm 6.58 \times 10^{-3}$
0.98	4.214×10^{-4}	4.382×10^{-4}	4.311×10^{-4}	6.172×10^{-4}	6.49×10^{-4}	3.614×10^{-3}	1.204×10^{-2}
	$\pm 2.02 \times 10^{-5}$	$\pm 2.00 \times 10^{-5}$	$\pm 1.99 \times 10^{-5}$	$\pm 2.25 \times 10^{-5}$	$\pm 2.20 \times 10^{-5}$	$\pm 1.53 \times 10^{-4}$	$\pm 1.016 \times 10^{-3}$
0.95	1.710×10^{-4}	1.895×10^{-4}	1.810×10^{-4}	4.143×10^{-4}	4.532×10^{-4}	2.322×10^{-3}	1.960×10^{-3}
	$\pm 9.11 \times 10^{-6}$	$\pm 8.67 \times 10^{-6}$	$\pm 8.49 \times 10^{-6}$	$\pm 8.48 \times 10^{-6}$	$\pm 1.11 \times 10^{-5}$	$\pm 1.51 \times 10^{-4}$	$\pm 1.85 \times 10^{-4}$
0.93	1.264×10^{-4}	1.465×10^{-4}	1.362×10^{-4}	3.921×10^{-4}	4.309×10^{-4}	1.638×10^{-3}	1.069×10^{-3}
	$\pm 5.59 \times 10^{-6}$	$\pm 4.76 \times 10^{-6}$	$\pm 4.70 \times 10^{-6}$	$\pm 1.85 \times 10^{-5}$	$\pm 2.07 \times 10^{-5}$	$\pm 1.067 \times 10^{-4}$	$\pm 1.022 \times 10^{-4}$
0.9	8.507×10^{-5}	1.060×10^{-4}	9.450×10^{-5}	3.382×10^{-4}	3.655×10^{-4}	1.008×10^{-3}	5.099×10^{-4}
	$\pm 1.43 \times 10^{-5}$	$\pm 1.44 \times 10^{-5}$	$\pm 1.35 \times 10^{-5}$	$\pm 5.19 \times 10^{-5}$	$\pm 5.73 \times 10^{-5}$	$\pm 2.04 \times 10^{-4}$	$\pm 1.15 \times 10^{-4}$

Table 47: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2E[\widehat{W \vert W > 0}], 4E[\widehat{W \vert W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $D/M/s$ model with $s = 400$ for actual delays in $(4E[\widehat{W\|W} > 0], 6E[\widehat{W\|W} > 0])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.459×10^{-3}	1.440×10^{-3}	1.432×10^{-3}	1.608×10^{-3}	1.635×10^{-3}	4.402×10^{-3}	1.981×10^{-1}
	$\pm 1.07 \times 10^{-4}$	$\pm 1.062 \times 10^{-4}$	$\pm 1.06 \times 10^{-4}$	$\pm 1.17 \times 10^{-4}$	$\pm 1.16 \times 10^{-4}$	$\pm 3.88 \times 10^{-4}$	$\pm 1.72 \times 10^{-2}$
0.98	7.847×10^{-4}	7.742×10^{-4}	7.667×10^{-4}	9.634×10^{-4}	9.909×10^{-4}	4.512×10^{-3}	5.460×10^{-2}
	$\pm 4.20 \times 10^{-5}$	$\pm 4.16 \times 10^{-5}$	$\pm 4.13 \times 10^{-5}$	$\pm 4.75 \times 10^{-5}$	$\pm 4.18 \times 10^{-5}$	$\pm 5.29 \times 10^{-4}$	$\pm 5.23 \times 10^{-4}$
0.95	3.415×10^{-4}	3.346×10^{-4}	3.238×10^{-4}	5.831×10^{-4}	6.308×10^{-4}	4.796×10^{-3}	8.892×10^{-3}
	$\pm 2.05 \times 10^{-5}$	$\pm 1.87 \times 10^{-5}$	$\pm 1.87 \times 10^{-5}$	$\pm 3.67 \times 10^{-5}$	$\pm 3.71 \times 10^{-5}$	$\pm 1.88 \times 10^{-4}$	$\pm 1.00 \times 10^{-3}$
0.93	2.609×10^{-4}	2.450×10^{-4}	2.340×10^{-4}	5.750×10^{-4}	6.412×10^{-4}	4.051×10^{-3}	4.525×10^{-3}
	$\pm 1.84 \times 10^{-5}$	$\pm 1.59 \times 10^{-5}$	$\pm 1.50 \times 10^{-5}$	$\pm 8.49 \times 10^{-5}$	$\pm 1.17 \times 10^{-4}$	$\pm 3.87 \times 10^{-4}$	$\pm 4.48 \times 10^{-4}$
0.9	2.475×10^{-4}	2.549×10^{-4}	2.275×10^{-4}	8.031×10^{-4}	9.401×10^{-4}	2.797×10^{-3}	2.101×10^{-3}
	$\pm 8.86 \times 10^{-5}$	$\pm 9.52 \times 10^{-5}$	$\pm 8.06 \times 10^{-5}$	$\pm 3.37 \times 10^{-4}$	$\pm 3.72 \times 10^{-4}$	$\pm 7.33 \times 10^{-4}$	$\pm 6.53 \times 10^{-4}$

Table 48: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

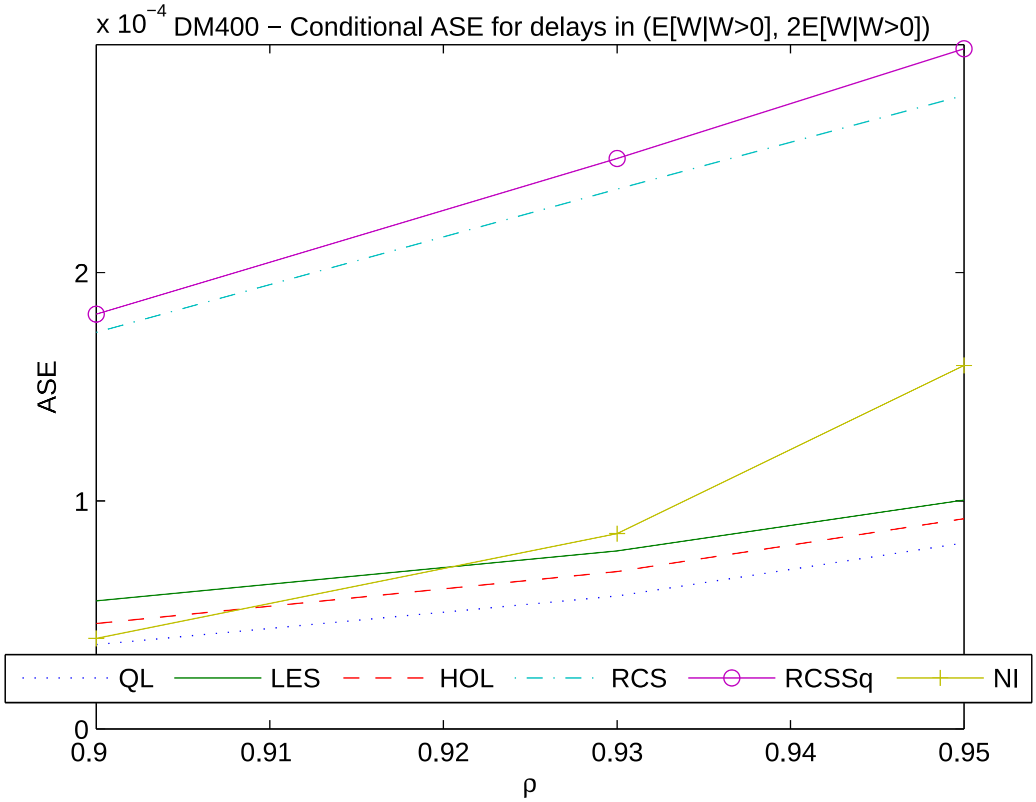

Figure 46: Conditional ASE for the alternative delay estimators in the $D/M/400$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

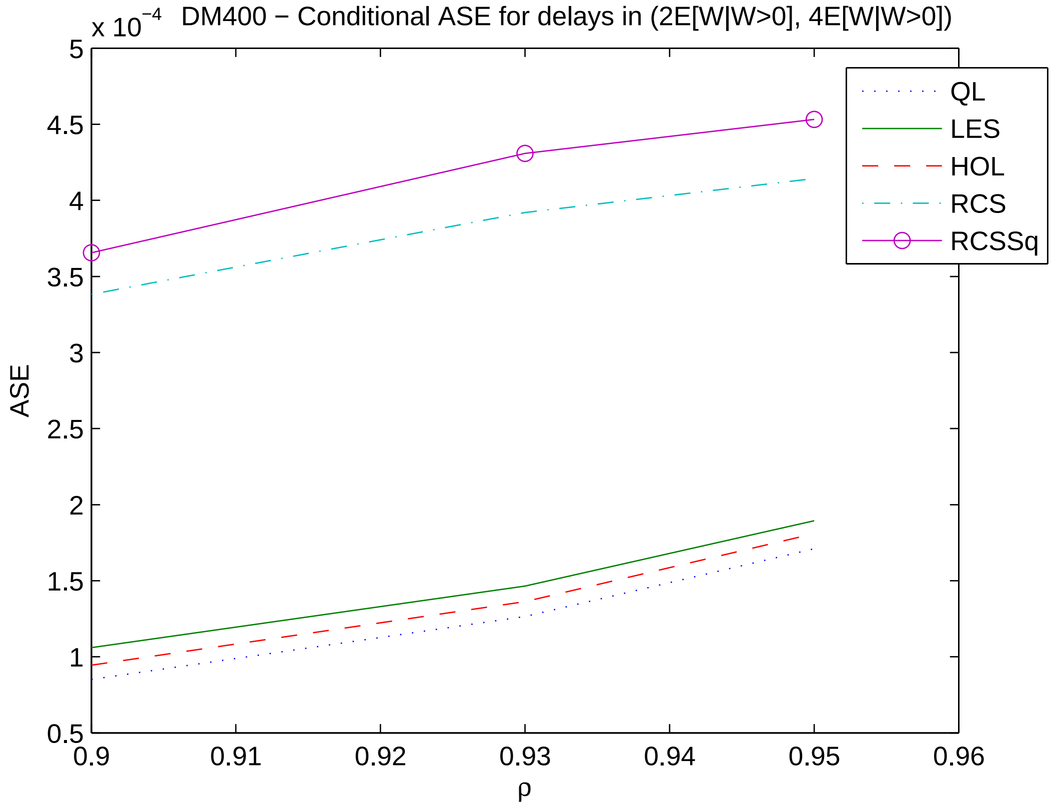

Figure 47: Conditional ASE for the alternative delay estimators in the $D/M/400$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

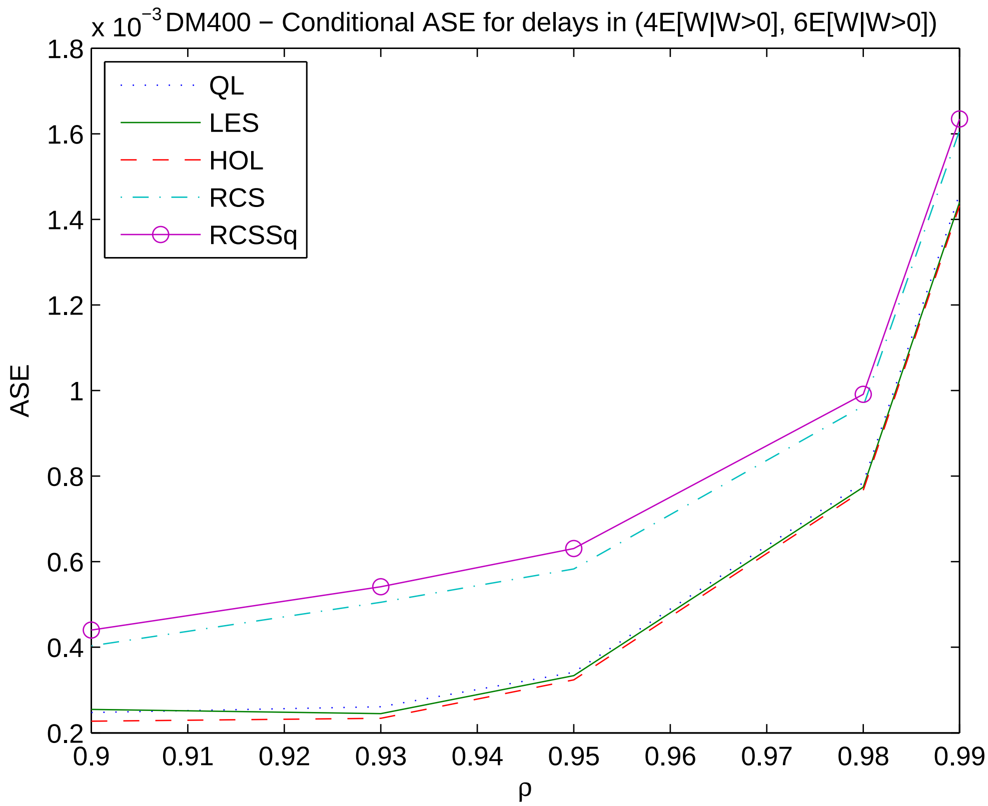

Figure 48: Conditional ASE for the alternative delay estimators in the $D/M/400$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

	Conditional ASE in the D/M/s model with $s = 900$ *for actual delays in* $(E[\widehat{W\|W} > 0], 2E[\widehat{W\|W} > 0])$							
ρ	$E[\widehat{W\|W} > 0]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	5.391×10^{-2}	8.339×10^{-5}	8.680×10^{-5}	8.547×10^{-5}	1.363×10^{-4}	1.445×10^{-4}	1.117×10^{-3}	7.251×10^{-4}
	$\pm 3.09 \times 10^{-3}$	$\pm 5.01 \times 10^{-6}$	$\pm 5.10 \times 10^{-6}$	$\pm 5.09 \times 10^{-6}$	$\pm 5.27 \times 10^{-6}$	$\pm 5.44 \times 10^{-6}$	$\pm 5.12 \times 10^{-5}$	$\pm 8.05 \times 10^{-5}$
0.98	2.782×10^{-2}	4.254×10^{-5}	4.604×10^{-5}	4.465×10^{-5}	9.869×10^{-5}	1.072×10^{-4}	7.203×10^{-4}	1.954×10^{-4}
	$\pm 1.23 \times 10^{-3}$	$\pm 2.09 \times 10^{-6}$	$\pm 2.17 \times 10^{-6}$	$\pm 2.15 \times 10^{-6}$	$\pm 2.57 \times 10^{-6}$	$\pm 2.63 \times 10^{-6}$	$\pm 4.00 \times 10^{-5}$	$\pm 1.58 \times 10^{-5}$
0.95	1.1972×10^{-2}	1.732×10^{-5}	2.104×10^{-5}	1.948×10^{-5}	7.079×10^{-5}	7.550×10^{-5}	2.458×10^{-4}	3.582×10^{-5}
	1.275×10^{-3}	1.853×10^{-6}	1.983×10^{-6}	1.979×10^{-6}	5.254×10^{-6}	6.201×10^{-6}	4.465×10^{-5}	7.230×10^{-6}
0.93	8.552×10^{-3}	1.214×10^{-5}	1.595×10^{-5}	1.419×10^{-5}	6.063×10^{-5}	6.352×10^{-5}	1.416×10^{-4}	2.002×10^{-5}
	$\pm 1.39 \times 10^{-3}$	$\pm 2.08 \times 10^{-6}$	2.27×10^{-6}	2.35×10^{-6}	$\pm 7.30 \times 10^{-6}$	$\pm 8.47 \times 10^{-6}$	$\pm 3.86 \times 10^{-5}$	6.15×10^{-6}

Table 49: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the D/M/s model with s = 900 for actual delays in $(2\widehat{E[W\|W>0]}, 4\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.99	1.640×10^{-4}	1.674×10^{-4}	1.660×10^{-4}	2.163×10^{-4}	2.248×10^{-4}	1.522×10^{-3}	9.410×10^{-3}
	$\pm 1.17 \times 10^{-5}$	$\pm 1.16 \times 10^{-5}$	$\pm 1.16 \times 10^{-5}$	$\pm 1.25 \times 10^{-5}$	$\pm 1.25 \times 10^{-5}$	$\pm 9.62 \times 10^{-5}$	$\pm 1.214 \times 10^{-3}$
0.98	8.442×10^{-5}	8.783×10^{-5}	8.641×10^{-5}	1.417×10^{-4}	1.514×10^{-4}	1.404×10^{-3}	2.486×10^{-3}
	$\pm 4.18 \times 10^{-6}$	$\pm 4.19 \times 10^{-6}$	$\pm 4.19 \times 10^{-6}$	$\pm 4.26 \times 10^{-6}$	$\pm 4.69 \times 10^{-6}$	$\pm 8.40 \times 10^{-5}$	$\pm 2.56 \times 10^{-4}$
0.95	3.684×10^{-5}	4.034×10^{-5}	3.862×10^{-5}	1.11378×10^{-4}	1.23493×10^{-4}	7.120×10^{-4}	4.66×10^{-4}
	$\pm 4.98 \times 10^{-6}$	$\pm 4.75 \times 10^{-6}$	$\pm 4.75 \times 10^{-6}$	$\pm 5.902 \times 10^{-6}$	$\pm 6.39 \times 10^{-6}$	$\pm 1.09 \times 10^{-4}$	$\pm 1.06 \times 10^{-4}$
0.93	2.605×10^{-5}	2.918×10^{-5}	2.734×10^{-5}	9.176×10^{-5}	9.681×10^{-5}	4.251×10^{-4}	2.152×10^{-4}
	$\pm 5.32 \times 10^{-6}$	$\pm 4.88 \times 10^{-6}$	$\pm 4.87 \times 10^{-6}$	$\pm 1.04 \times 10^{-5}$	$\pm 9.56 \times 10^{-6}$	1.06×10^{-4}	6.52×10^{-5}

Table 50: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W|W>0]}, 4\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

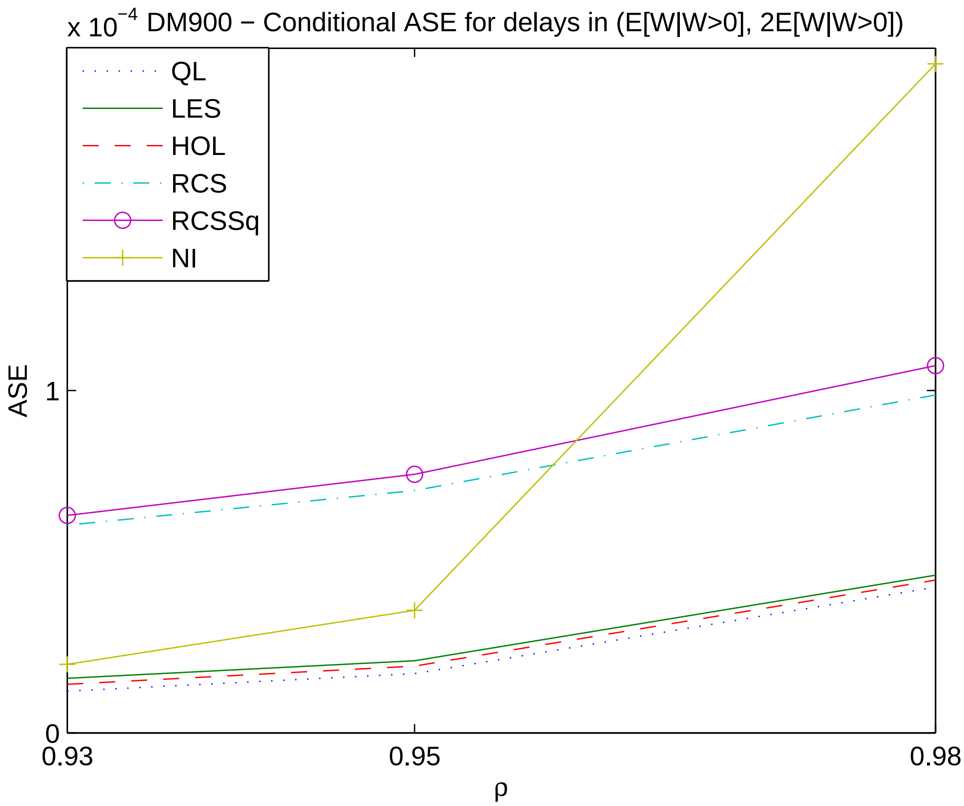

Figure 49: Conditional ASE for the alternative delay estimators in the $D/M/900$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

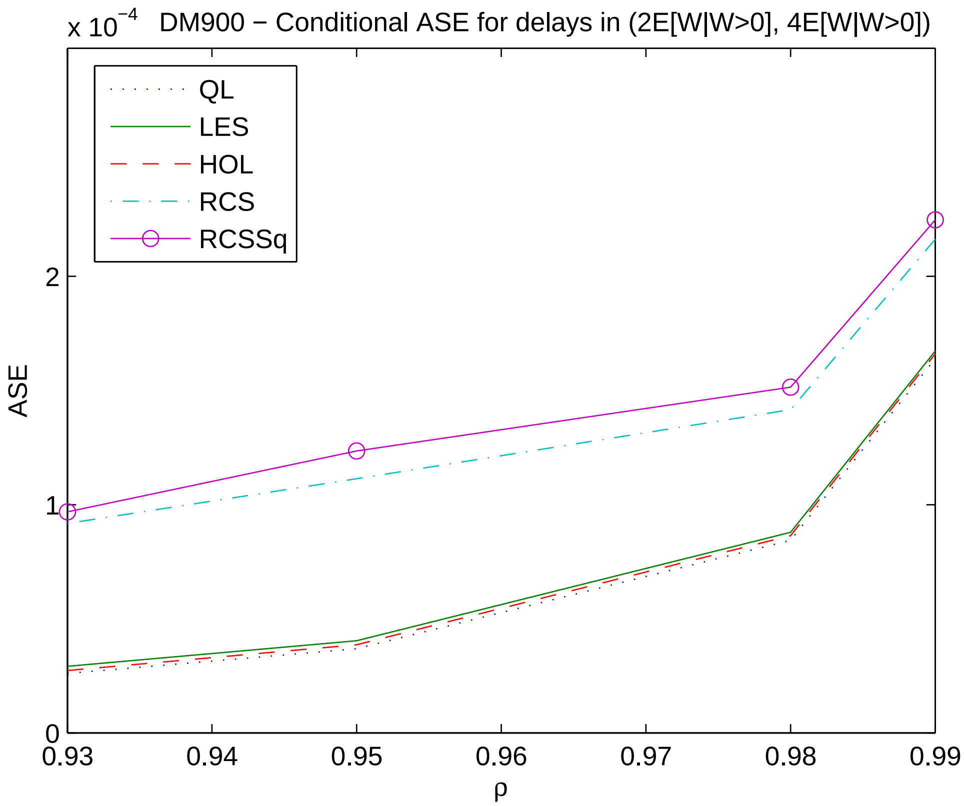

Figure 50: Conditional ASE for the alternative delay estimators in the $D/M/900$ model for actual delays in $(2E[\widehat{W|W > 0}], 4E[\widehat{W|W > 0}])$, as a function of the traffic intensity ρ

Conditional ASE in the $H_2/M/s$ model with $s = 100$ for actual delays in $(E[\widehat{W\|W > 0}], 2E[\widehat{W\|W > 0}])$								
ρ	$E[\widehat{W\|W > 0}]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.283	1.804×10^{-2}	8.789×10^{-2}	8.740×10^{-2}	9.381×10^{-2}	9.481×10^{-2}	1.381×10^{-1}	4.222×10^{-1}
	$\pm 1.13 \times 10^{-1}$	$\pm 1.54 \times 10^{-3}$	$\pm 7.78 \times 10^{-3}$	$\pm 7.780 \times 10^{-3}$	$\pm 7.80 \times 10^{-3}$	7.78×10^{-3}	$\pm 8.43 \times 10^{-3}$	$\pm 7.18 \times 10^{-2}$
0.95	4.838×10^{-1}	6.728×10^{-3}	3.123×10^{-1}	3.0739×10^{-2}	3.720×10^{-2}	3.819×10^{-2}	7.849×10^{-2}	5.927×10^{-2}
	$\pm 1.54 \times 10^{-2}$	$\pm 2.05 \times 10^{-4}$	$\pm 8.15 \times 10^{-4}$	$\pm 8.15 \times 10^{-4}$	$\pm 7.70 \times 10^{-4}$	$\pm 7.69 \times 10^{-4}$	$\pm 1.31 \times 10^{-3}$	$\pm 3.74 \times 10^{-3}$
0.93	3.430×10^{-1}	4.750×10^{-3}	2.137×10^{-2}	2.089×10^{-2}	2.729×10^{-2}	2.826×10^{-2}	6.222×10^{-2}	2.982×10^{-2}
	$\pm 8.45 \times 10^{-3}$	$\pm 1.32 \times 10^{-4}$	$\pm 5.72 \times 10^{-4}$	$\pm 5.72 \times 10^{-4}$	$\pm 6.39 \times 10^{-4}$	$\pm 6.61 \times 10^{-4}$	$\pm 1.34 \times 10^{-3}$	$\pm 1.43 \times 10^{-3}$
0.9	2.3620×10^{-1}	3.226×10^{-3}	1.399×10^{-2}	1.351×10^{-2}	1.987×10^{-2}	2.075×10^{-2}	4.563×10^{-2}	1.411×10^{-2}
	$\pm 3.83 \times 10^{-3}$	$\pm 6.53 \times 10^{-5}$	$\pm 2.96 \times 10^{-4}$	$\pm 2.89 \times 10^{-4}$	$\pm 3.77 \times 10^{-4}$	$\pm 3.88 \times 10^{-4}$	$\pm 8.59 \times 10^{-4}$	$\pm 4.22 \times 10^{-4}$

Table 51: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE *in the* $H_2/M/s$ *model with* $s = 100$ *for actual delays in* $(2\widehat{E[W\|W>0]}, 4\widehat{E[W\|W>0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	3.596×10^{-2}	1.744×10^{-1}	1.739×10^{-1}	1.807×10^{-1}	1.818×10^{-1}	2.270×10^{-1}	5.513
	$\pm 3.88 \times 10^{-3}$	$\pm 1.72 \times 10^{-2}$	$\pm 1.719 \times 10^{-2}$	$\pm 1.73 \times 10^{-2}$	$\pm 1.73 \times 10^{-2}$	$\pm 1.80 \times 10^{-2}$	1.02
0.95	1.296×10^{-2}	6.052×10^{-2}	6.000×10^{-2}	6.692×10^{-2}	6.800×10^{-2}	1.170×10^{-1}	7.268×10^{-1}
	$\pm 4.66 \times 10^{-4}$	$\pm 1.642 \times 10^{-3}$	$\pm 1.65 \times 10^{-3}$	$\pm 1.53 \times 10^{-3}$	$\pm 1.57 \times 10^{-3}$	$\pm 2.44 \times 10^{-3}$	$\pm 4.55 \times 10^{-2}$
0.93	9.234×10^{-3}	4.210×10^{-2}	4.156×10^{-2}	4.854×10^{-2}	4.963×10^{-2}	10.00×10^{-2}	3.640×10^{-1}
	$\pm 2.52 \times 10^{-4}$	$\pm 1.08 \times 10^{-3}$	$\pm 1.08 \times 10^{-3}$	$\pm 1.12 \times 10^{-3}$	$\pm 1.13 \times 10^{-3}$	$\pm 1.98 \times 10^{-3}$	$\pm 1.66 \times 10^{-2}$
0.9	6.391×10^{-3}	2.788×10^{-2}	2.734×10^{-2}	3.449×10^{-2}	3.565×10^{-2}	8.357×10^{-2}	1.744×10^{-1}
	$\pm 1.22 \times 10^{-4}$	$\pm 5.92 \times 10^{-4}$	$\pm 5.89 \times 10^{-4}$	$\pm 7.37 \times 10^{-4}$	$\pm 7.65 \times 10^{-4}$	$\pm 1.81 \times 10^{-3}$	$\pm 6.28 \times 10^{-3}$

Table 52: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W|W>0]}, 4\widehat{E[W|W>0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ *model with* $s = 100$ *for actual delays in* $(4E[\widehat{W|W > 0}], 6E[\widehat{W|W > 0}])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	5.359×10^{-2}	2.836×10^{-1}	2.831×10^{-1}	2.894×10^{-1}	2.904×10^{-1}	3.348×10^{-1}	21.34
	$\pm 6.15 \times 10^{-3}$	$\pm 4.72 \times 10^{-2}$	$\pm 4.71 \times 10^{-2}$	$\pm 4.81 \times 10^{-2}$	$\pm 4.83 \times 10^{-2}$	$\pm 5.16 \times 10^{-2}$	± 4.21
0.95	2.376×10^{-2}	1.122×10^{-1}	1.116×10^{-1}	1.190×10^{-1}	1.201×10^{-1}	1.706×10^{-1}	3.305
	$\pm 1.64 \times 10^{-3}$	$\pm 7.04 \times 10^{-3}$	$\pm 6.96 \times 10^{-3}$	$\pm 7.30 \times 10^{-3}$	$\pm 7.31 \times 10^{-3}$	$\pm 9.69 \times 10^{-3}$	2.26×10^{-1}
0.93	1.696×10^{-2}	7.824×10^{-2}	7.7659×10^{-2}	8.548×10^{-2}	8.661×10^{-2}	1.466×10^{-1}	1.654
	$\pm 8.96 \times 10^{-4}$	$\pm 4.30 \times 10^{-3}$	$\pm 4.26 \times 10^{-3}$	$\pm 4.66 \times 10^{-3}$	$\pm 4.66 \times 10^{-3}$	$\pm 7.65 \times 10^{-3}$	$\pm 7.37 \times 10^{-2}$
0.90	1.184×10^{-2}	5.373×10^{-2}	5.307×10^{-2}	6.140×10^{-2}	6.276×10^{-2}	1.291×10^{-1}	7.745×10^{-1}
	$\pm 3.24 \times 10^{-4}$	$\pm 2.37 \times 10^{-3}$	$\pm 2.36 \times 10^{-3}$	$\pm 3.00 \times 10^{-3}$	3.082×10^{-3}	$\pm 5.48 \times 10^{-3}$	$\pm 2.59 \times 10^{-2}$

Table 53: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4E[\widehat{W|W > 0}], 6E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ model with $s = 100$ for actual delays $> 6E[\widehat{W \vert W > 0}]$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	3.210×10^{-2}	1.741×10^{-1}	1.734×10^{-1}	1.821×10^{-1}	1.832×10^{-1}	2.454×10^{-1}	8.023
	$\pm 7.23 \times 10^{-3}$	$\pm 3.85 \times 10^{-2}$	$\pm 3.83 \times 10^{-2}$	$\pm 3.97 \times 10^{-2}$	$\pm 3.97 \times 10^{-2}$	$\pm 5.08 \times 10^{-2}$	± 1.11
0.93	2.661×10^{-2}	1.270×10^{-1}	1.263×10^{-1}	1.340×10^{-1}	1.353×10^{-1}	1.962×10^{-1}	4.280
	$\pm 1.97 \times 10^{-3}$	$\pm 1.44 \times 10^{-2}$	$\pm 1.43 \times 10^{-2}$	$\pm 1.56 \times 10^{-2}$	$\pm 1.59 \times 10^{-2}$	$\pm 2.54 \times 10^{-2}$	$\pm 4.31 \times 10^{-1}$
0.9	1.884×10^{-2}	9.237×10^{-2}	9.158×10^{-2}	1.019×10^{-1}	1.037×10^{-1}	1.849×10^{-1}	1.977
	$\pm 2.73 \times 10^{-3}$	$\pm 7.55 \times 10^{-3}$	$\pm 7.55 \times 10^{-3}$	$\pm 7.60 \times 10^{-3}$	$\pm 7.64 \times 10^{-3}$	$\pm 1.30 \times 10^{-2}$	$\pm 1.11 \times 10^{-1}$

Table 54: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) for delays larger than $6E[\widehat{W \vert W > 0}]$). Each estimate is shown with the half width of the 95 percent confidence interval.

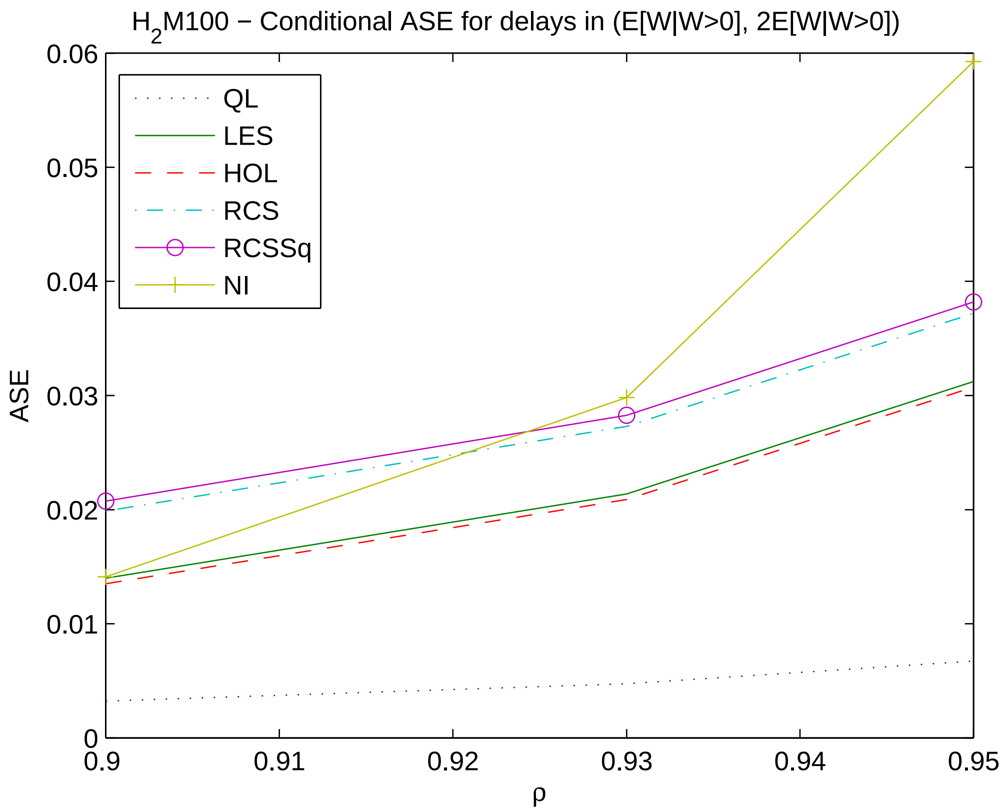

Figure 51: Conditional ASE for the alternative delay estimators in the $H_2/M/100$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

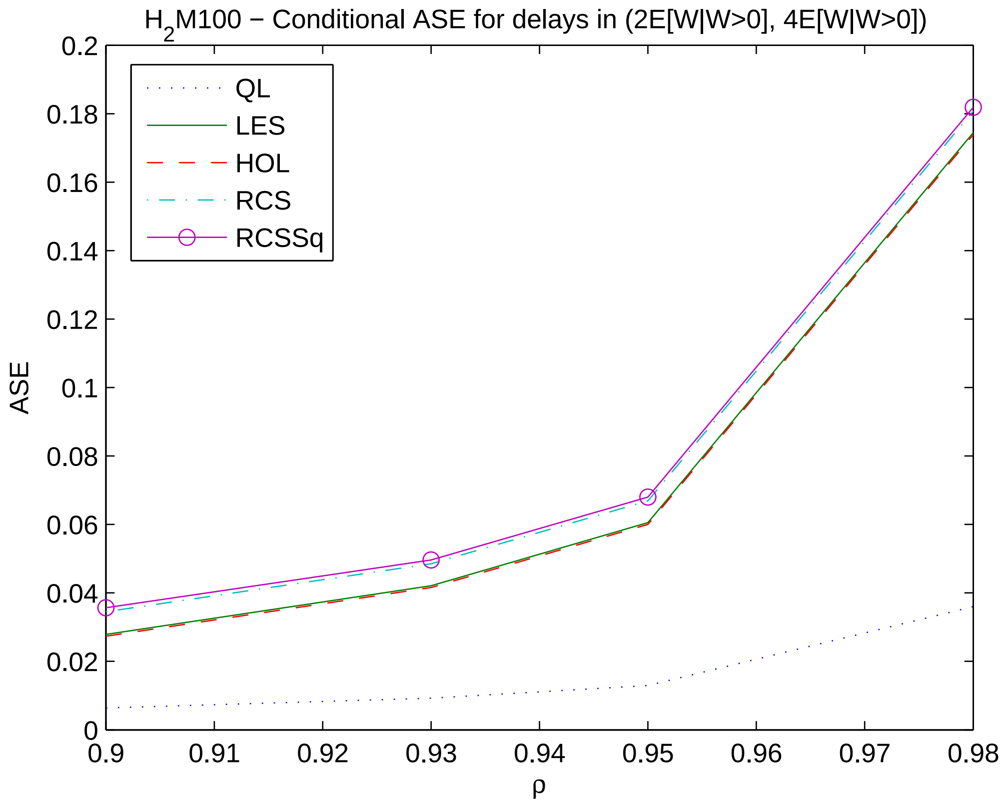

Figure 52: Conditional ASE for the alternative delay estimators in the $H_2/M/900$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

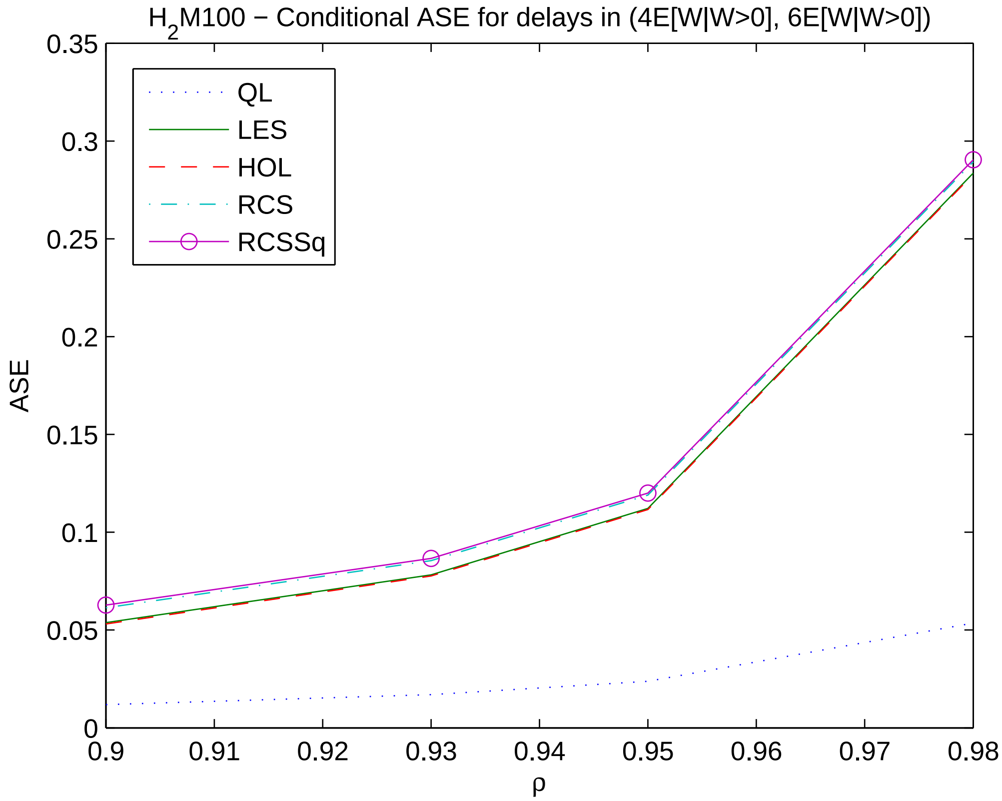

Figure 53: Conditional ASE for the alternative delay estimators in the $H_2/M/100$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

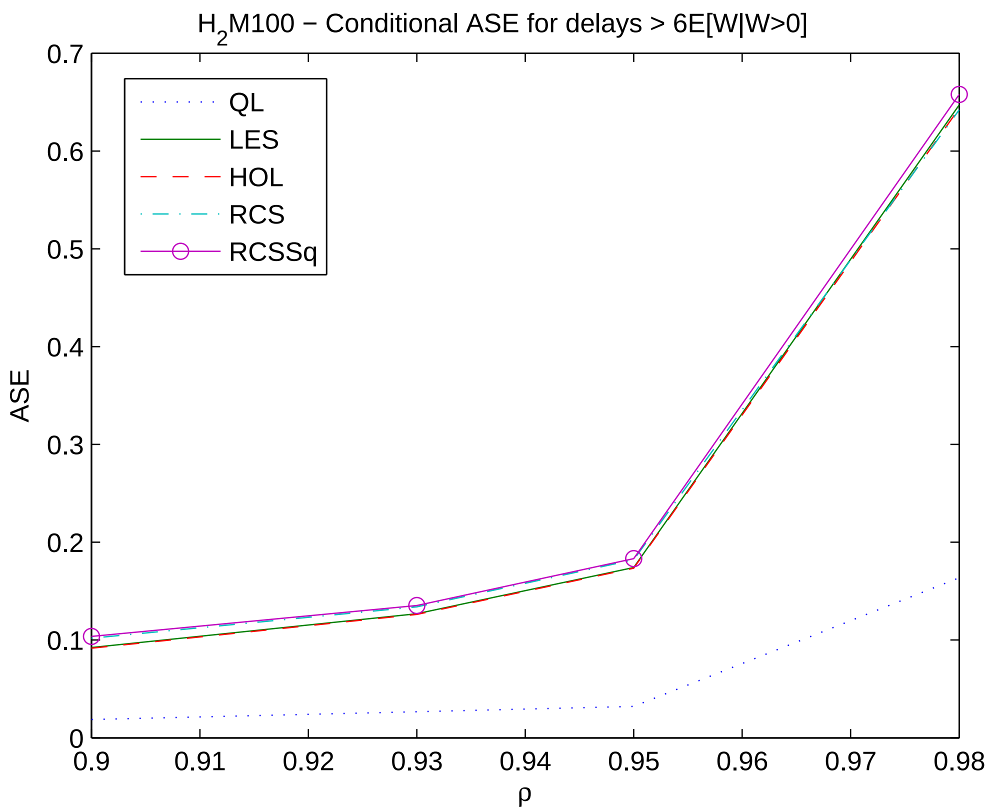

Figure 54: Conditional ASE for the alternative delay estimators in the $H_2/M/100$ model for actual delays larger than $6E[\widehat{W|W} > 0]$, as a function of the traffic intensity ρ

Conditional ASE in the $H_2/M/s$ model with $s = 400$ for actual delays in $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$

ρ	$E[\widehat{W\|W > 0}]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	2.978×10^{-1}	1.056×10^{-3}	5.122×10^{-3}	5.091×10^{-3}	5.882×10^{-3}	6.01×10^{-3}	1.759×10^{-2}	2.273×10^{-2}
	$\pm 1.906 \times 10^{-2}$	$\pm 7.32 \times 10^{-5}$	$\pm 3.22 \times 10^{-4}$	$\pm 3.21 \times 10^{-4}$	$\pm 3.28 \times 10^{-4}$	$\pm 3.26 \times 10^{-4}$	$\pm 7.86 \times 10^{-4}$	$\pm 2.87 \times 10^{-3}$
0.95	1.218×10^{-1}	4.293×10^{-4}	1.969×10^{-3}	1.939×10^{-3}	2.738×10^{-3}	2.868×10^{-3}	1.068×10^{-2}	3.715×10^{-3}
	$\pm 4.93 \times 10^{-3}$	$\pm 1.78 \times 10^{-5}$	$\pm 7.66 \times 10^{-5}$	$\pm 7.67 \times 10^{-5}$	$\pm 8.65 \times 10^{-5}$	$\pm 8.57 \times 10^{-5}$	$\pm 4.10 \times 10^{-4}$	$\pm 2.65 \times 10^{-4}$
0.93	8.611×10^{-2}	2.977×10^{-4}	1.339×10^{-3}	1.309×10^{-3}	2.083×10^{-3}	2.205×10^{-3}	7.568×10^{-3}	1.882×10^{-3}
	$\pm 2.55 \times 10^{-3}$	$\pm 9.76 \times 10^{-6}$	$\pm 5.05 \times 10^{-5}$	$\pm 5.06 \times 10^{-5}$	$\pm 5.82 \times 10^{-5}$	$\pm 6.36 \times 10^{-5}$	$\pm 3.18 \times 10^{-4}$	$\pm 1.14 \times 10^{-4}$
0.9	5.845×10^{-2}	1.981×10^{-4}	8.575×10^{-4}	8.283×10^{-4}	1.562×10^{-3}	1.657×10^{-3}	4.736×10^{-3}	8.672×10^{-4}
	$\pm 1.58 \times 10^{-3}$	$\pm 5.65 \times 10^{-6}$	$\pm 3.11 \times 10^{-5}$	$\pm 3.07 \times 10^{-5}$	$\pm 4.12 \times 10^{-5}$	$\pm 4.76 \times 10^{-5}$	$\pm 2.15 \times 10^{-4}$	$\pm 3.99 \times 10^{-5}$

Table 55: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W > 0}], 2E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ *model with* $s = 400$ *for actual delays in* $(2\widehat{E[W\|W > 0]}, 4\widehat{E[W\|W > 0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	2.012×10^{-3}	9.638×10^{-3}	9.605×10^{-3}	1.041×10^{-2}	1.054×10^{-2}	2.333×10^{-2}	2.787×10^{-1}
	$\pm 1.35 \times 10^{-4}$	$\pm 8.16 \times 10^{-4}$	$\pm 8.15 \times 10^{-4}$	$\pm 8.47 \times 10^{-4}$	$\pm 8.53 \times 10^{-4}$	$\pm 1.37 \times 10^{-3}$	$\pm 3.99 \times 10^{-2}$
0.95	8.336×10^{-4}	3.811×10^{-3}	3.778×10^{-3}	4.616×10^{-3}	4.752×10^{-3}	1.812×10^{-2}	4.616×10^{-2}
	$\pm 2.97 \times 10^{-5}$	$\pm 1.99 \times 10^{-4}$	$\pm 1.997 \times 10^{-4}$	$\pm 2.105 \times 10^{-4}$	$\pm 2.089 \times 10^{-4}$	$\pm 5.503 \times 10^{-4}$	$\pm 3.516 \times 10^{-3}$
0.93	5.766×10^{-4}	2.626×10^{-3}	2.594×10^{-3}	3.450×10^{-3}	3.595×10^{-3}	1.553×10^{-2}	2.271×10^{-3}
	$\pm 1.97 \times 10^{-5}$	$\pm 1.09 \times 10^{-4}$	$\pm 1.09 \times 10^{-4}$	$\pm 1.17 \times 10^{-4}$	$\pm 1.24 \times 10^{-4}$	$\pm 4.58 \times 10^{-4}$	$\pm 1.45 \times 10^{-3}$
0.9	3.927×10^{-4}	1.732×10^{-3}	1.699×10^{-3}	2.596×10^{-3}	2.74×10^{-3}	1.159×10^{-2}	1.07×10^{-2}
	$\pm 1.51 \times 10^{-5}$	$\pm 7.59 \times 10^{-5}$	$\pm 7.61 \times 10^{-5}$	$\pm 1.022 \times 10^{-4}$	$\pm 1.025 \times 10^{-4}$	$\pm 4.76 \times 10^{-4}$	$\pm 5.48 \times 10^{-4}$

Table 56: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2\widehat{E[W|W > 0]}, 4\widehat{E[W|W > 0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ model with $s = 400$ for actual delays in $(4E[\widehat{W\|W > 0}], 6E[\widehat{W\|W > 0}])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	3.238×10^{-3}	1.661×10^{-2}	1.659×10^{-2}	1.738×10^{-2}	1.751×10^{-2}	3.017×10^{-2}	1.193
$\pm 3.68 \times 10^{-4}$	$\pm 1.71 \times 10^{-3}$	$\pm 1.71 \times 10^{-3}$	$\pm 1.76 \times 10^{-3}$	$\pm 1.76 \times 10^{-3}$	$\pm 2.80 \times 10^{-3}$	$\pm 1.73 \times 10^{-1}$	
0.95	1.540×10^{-3}	6.717×10^{-3}	6.683×10^{-3}	7.603×10^{-3}	7.773×10^{-3}	2.514×10^{-2}	2.064×10^{-1}
	$\pm 8.31 \times 10^{-5}$	$\pm 7.02 \times 10^{-4}$	$\pm 7.00 \times 10^{-4}$	$\pm 8.77 \times 10^{-4}$	$\pm 8.95 \times 10^{-4}$	$\pm 2.64 \times 10^{-3}$	$\pm 1.854 \times 10^{-2}$
0.93	1.073×10^{-3}	5.139×10^{-3}	5.099×10^{-3}	6.217×10^{-3}	6.421×10^{-3}	2.722×10^{-2}	1.014×10^{-1}
	$\pm 1.21 \times 10^{-4}$	$\pm 3.462 \times 10^{-4}$	$\pm 3.45 \times 10^{-4}$	$\pm 3.54 \times 10^{-4}$	$\pm 3.75 \times 10^{-4}$	1.650×10^{-3}	7.38×10^{-3}
0.9	7.412×10^{-4}	3.253×10^{-3}	3.218×10^{-3}	4.174×10^{-3}	4.330×10^{-3}	2.287×10^{-2}	4.737×10^{-2}
	$\pm 6.18 \times 10^{-5}$	$\pm 2.93 \times 10^{-4}$	$\pm 2.91 \times 10^{-4}$	$\pm 3.63 \times 10^{-4}$	$\pm 4.09 \times 10^{-4}$	$\pm 1.788 \times 10^{-3}$	$\pm 3.61 \times 10^{-3}$

Table 57: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4E[\widehat{W|W > 0}], 6E[\widehat{W|W > 0}])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ model with $s=100$ for actual delays $> 6\widehat{E[W\|W>0]}$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.95	2.162×10^{-3}	1.259×10^{-2}	1.256×10^{-2}	1.345×10^{-2}	1.355×10^{-2}	3.872×10^{-2}	4.950×10^{-1}
	$\pm 4.97 \times 10^{-4}$	$\pm 4.00 \times 10^{-3}$	$\pm 3.99 \times 10^{-3}$	$\pm 4.42 \times 10^{-3}$	$\pm 4.48 \times 10^{-3}$	$\pm 1.88 \times 10^{-2}$	$\pm 1.03 \times 10^{-1}$
0.93	1.516×10^{-3}	7.045×10^{-3}	7.00×10^{-3}	7.791×10^{-3}	8.022×10^{-3}	3.952×10^{-2}	2.461×10^{-1}
	$\pm 4.30 \times 10^{-4}$	$\pm 1.22 \times 10^{-3}$	$\pm 1.22 \times 10^{-3}$	$\pm 1.47 \times 10^{-3}$	$\pm 1.56 \times 10^{-3}$	$\pm 1.01 \times 10^{-2}$	5.43×10^{-2}
0.9	1.413×10^{-3}	5.797×10^{-3}	5.745×10^{-3}	7.135×10^{-3}	7.381×10^{-3}	3.704×10^{-2}	1.123×10^{-1}
	$\pm 2.90 \times 10^{-4}$	1.63×10^{-3}	$\pm 1.62 \times 10^{-3}$	$\pm 1.41 \times 10^{-3}$	$\pm 1.51 \times 10^{-3}$	$\pm 7.61 \times 10^{-3}$	$\pm 1.18 \times 10^{-2}$

Table 58: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) for delays larger than $6\widehat{E[W|W > 0]}$). Each estimate is shown with the half width of the 95 percent confidence interval.

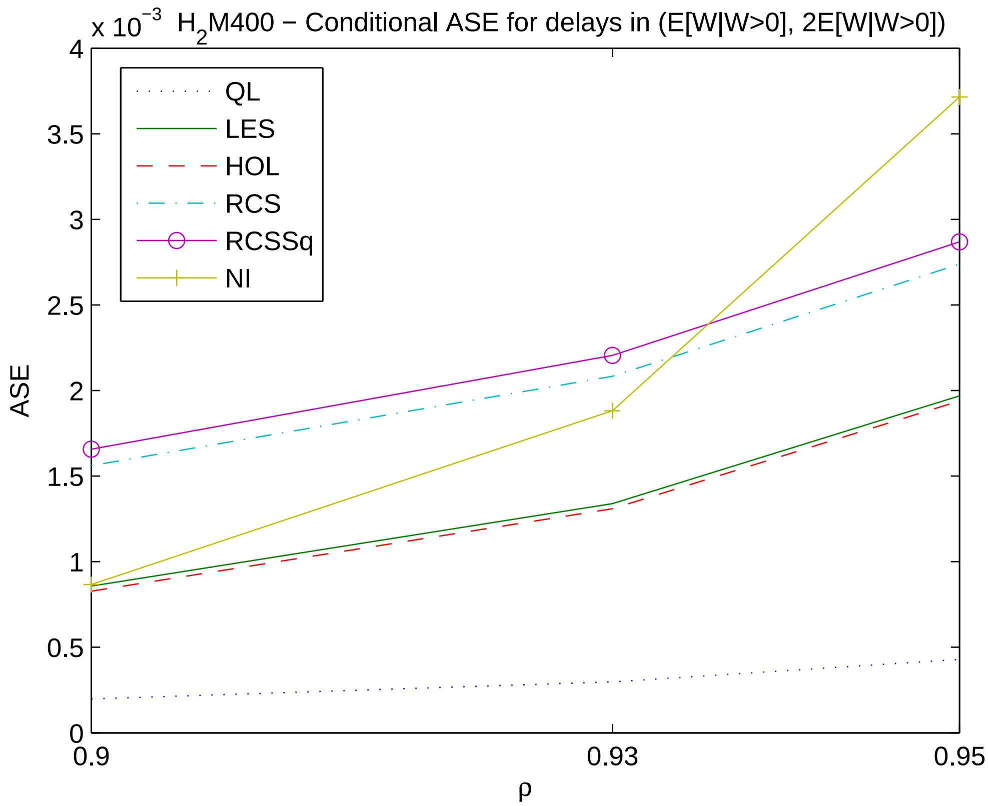

Figure 55: Conditional ASE for the alternative delay estimators in the $H_2/M/400$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

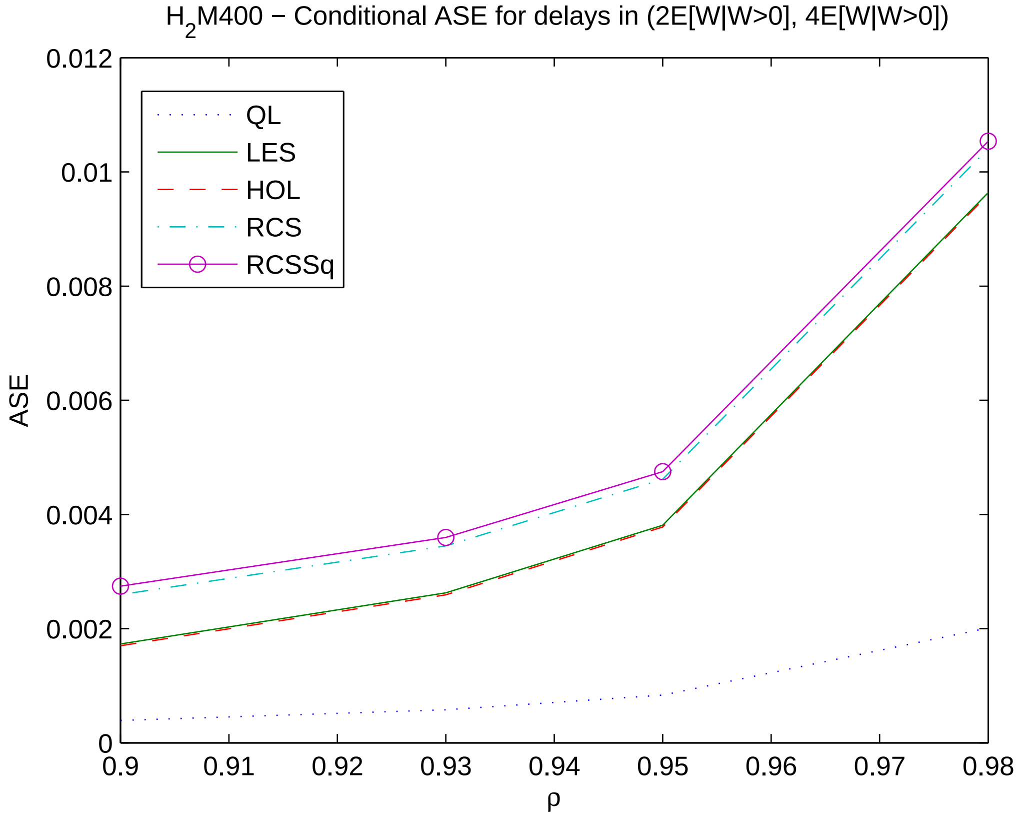

Figure 56: Conditional ASE for the alternative delay estimators in the $H_2/M/400$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

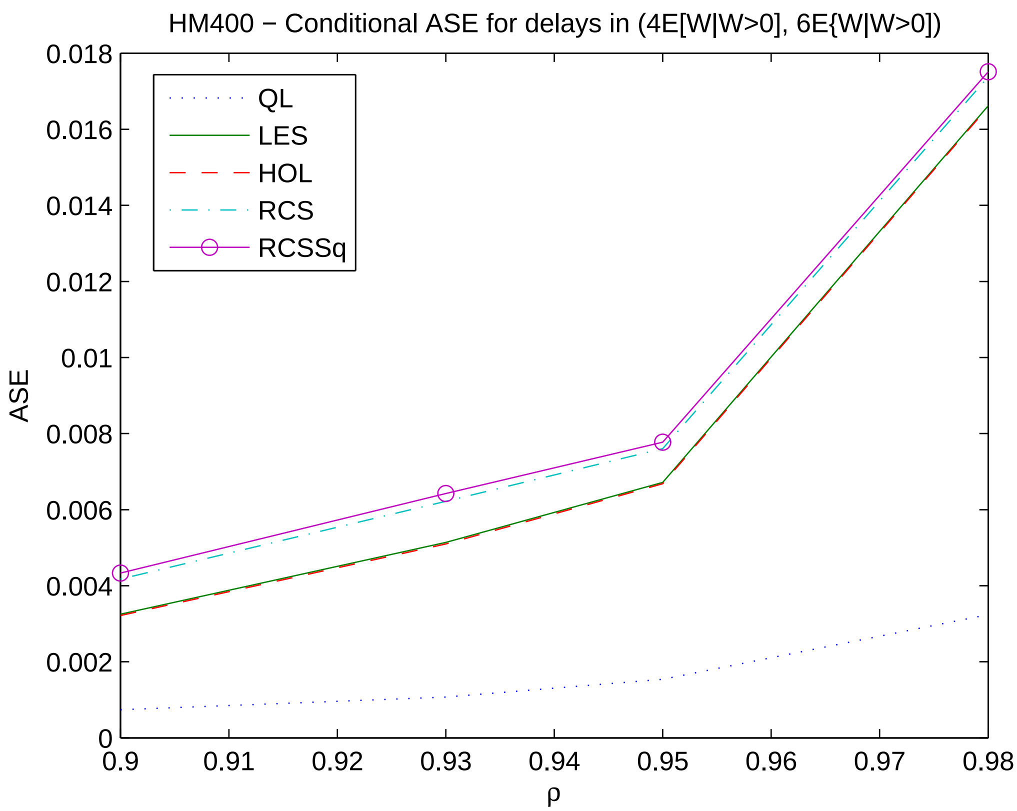

Figure 57: Conditional ASE for the alternative delay estimators in the $H_2/M/400$ model for actual delays in $(4E[\widehat{W|W} > 0], 6E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

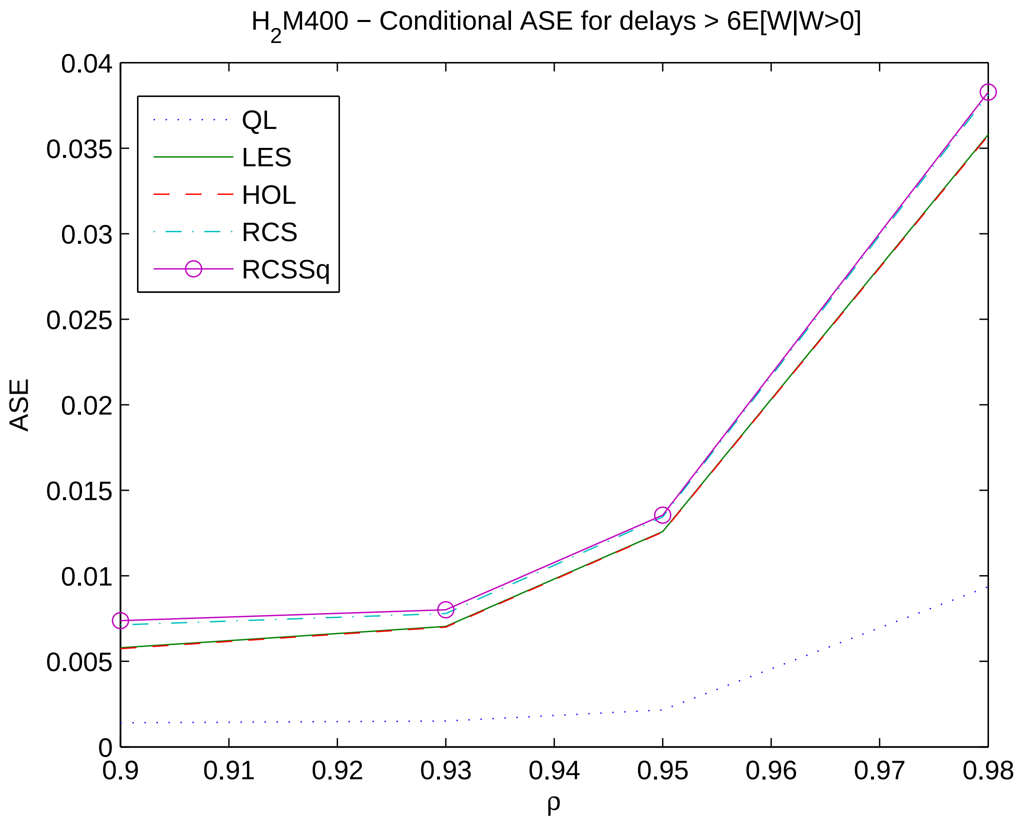

Figure 58: Conditional ASE for the alternative delay estimators in the $H_2/M/400$ model for actual delays larger than $6E[\widehat{W|W>0}]$, as a function of the traffic intensity ρ

	Conditional ASE in the $H_2/M/s$ model with $s = 900$ for actual delays in $(E[\widehat{W\|W} > 0], 2E[\widehat{W\|W} > 0])$							
ρ	$E[\widehat{W\|W} > 0]$	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS - \sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	1.3919×10^{-1}	2.174×10^{-4}	1.059×10^{-3}	1.053×10^{-3}	1.287×10^{-3}	1.325×10^{-3}	6.215×10^{-3}	4.946×10^{-3}
	$\pm 1.45 \times 10^{-2}$	$\pm 2.42 \times 10^{-5}$	$\pm 1.144 \times 10^{-4}$	1.14×10^{-4}	$\pm 1.14 \times 10^{-4}$	$\pm 1.14 \times 10^{-4}$	$\pm 3.57 \times 10^{-4}$	$\pm 1.04 \times 10^{-3}$
0.95	5.322×10^{-2}	8.219×10^{-5}	3.820×10^{-4}	3.760×10^{-4}	6.045×10^{-4}	6.409×10^{-4}	3.003×10^{-3}	7.254×10^{-4}
	$\pm 3.28 \times 10^{-3}$	$\pm 3.28 \times 10^{-3}$	$\pm 4.90 \times 10^{-6}$	$\pm 2.38 \times 10^{-5}$	$\pm 2.39 \times 10^{-5}$	$\pm 2.50 \times 10^{-5}$	$\pm 2.55 \times 10^{-5}$	$\pm 2.27 \times 10^{-4}$
0.93	3.811×10^{-2}	5.850×10^{-5}	2.620×10^{-4}	2.561×10^{-4}	4.827×10^{-4}	5.149×10^{-4}	2.003×10^{-3}	3.676×10^{-4}
	$\pm 1.92 \times 10^{-3}$	$\pm 3.12 \times 10^{-6}$	$\pm 1.42 \times 10^{-5}$	$\pm 1.41 \times 10^{-5}$	$\pm 1.69 \times 10^{-5}$	$\pm 1.87 \times 10^{-5}$	$\pm 1.58 \times 10^{-4}$	$\pm 3.75 \times 10^{-5}$

Table 59: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ model with $s = 900$ for actual delays in $(2E[\widehat{W\|W} > 0], 4E[\widehat{W\|W} > 0])$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	4.139×10^{-4}	1.982×10^{-3}	1.976×10^{-3}	2.208×10^{-3}	2.246132×10^{-3}	8.250×10^{-3}	6.160×10^{-2}
	$\pm 4.88 \times 10^{-5}$	$\pm 2.13 \times 10^{-4}$	$\pm 2.13 \times 10^{-4}$	$\pm 2.15 \times 10^{-4}$	$\pm 2.15 \times 10^{-4}$	$\pm 3.31 \times 10^{-4}$	$\pm 1.27 \times 10^{-2}$
0.95	1.591×10^{-4}	7.360×10^{-4}	7.299×10^{-4}	9.746×10^{-4}	1.015×10^{-3}	6.091×10^{-3}	8.771×10^{-3}
	$\pm 8.59 \times 10^{-5}$	$\pm 1.04 \times 10^{-5}$	$\pm 5.20 \times 10^{-5}$	$\pm 5.19 \times 10^{-5}$	$\pm 5.43 \times 10^{-5}$	$\pm 5.55 \times 10^{-5}$	$\pm 3.77 \times 10^{-4}$
0.93	1.137×10^{-4}	5.187×10^{-4}	5.121×10^{-4}	7.738×10^{-4}	8.171×10^{-4}	4.944×10^{-3}	4.500×10^{-3}
	$\pm 5.54 \times 10^{-6}$	$\pm 2.79 \times 10^{-6}$	$\pm 2.78 \times 10^{-5}$	$\pm 3.38 \times 10^{-5}$	$\pm 3.42 \times 10^{-5}$	$\pm 3.25 \times 10^{-4}$	$\pm 4.05 \times 10^{-4}$

Table 60: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$. Each estimate is shown with the half width of the 95 percent confidence interval.

Conditional ASE in the $H_2/M/s$ model with $s = 900$ for actual delays in $(4\widehat{E[W\|W > 0]}, 6\widehat{E[W\|W > 0]})$							
ρ	$ASE(QL)$	$ASE(LES)$	$ASE(HOL)$	$ASE(RCS)$	$ASE(RCS-\sqrt{s})$	$ASE(LCS)$	$ASE(NI)$
0.98	7.156×10^{-4}	3.564×10^{-3}	3.558×10^{-3}	3.788×10^{-3}	3.835×10^{-3}	9.868×10^{-3}	2.560×10^{-1}
	$\pm 9.46 \times 10^{-5}$	$\pm 4.31 \times 10^{-4}$	$\pm 4.31 \times 10^{-4}$	$\pm 4.34 \times 10^{-4}$	$\pm 4.37 \times 10^{-4}$	$\pm 8.63 \times 10^{-4}$	$\pm 4.10 \times 10^{-2}$
0.95	2.657×10^{-4}	1.349×10^{-3}	1.343×10^{-3}	1.619×10^{-3}	1.668×10^{-3}	1.091×10^{-2}	3.821×10^{-2}
	$\pm 1.06 \times 10^{-3}$	$\pm 2.47 \times 10^{-5}$	$\pm 1.31 \times 10^{-4}$	$\pm 1.30 \times 10^{-4}$	1.59×10^{-4}	1.640×10^{-4}	1.523×10^{-3}
0.93	2.289×10^{-4}	9.915×10^{-4}	9.836×10^{-4}	1.288×10^{-3}	1.343×10^{-3}	1.014×10^{-2}	2.018×10^{-2}
	$\pm 1.73 \times 10^{-5}$	$\pm 9.27 \times 10^{-5}$	$\pm 9.13 \times 10^{-5}$	1.08×10^{-4}	1.10×10^{-4}	$\pm 1.37 \times 10^{-3}$	$\pm 1.96 \times 10^{-2}$

Table 61: A comparison of the efficiency of different real-time delay estimators conditional on the level of delay observed for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the conditional average squared error (ASE) in the interval $(4\widehat{E[W|W > 0]}, 6\widehat{E[W|W > 0]})$. Each estimate is shown with the half width of the 95 percent confidence interval.

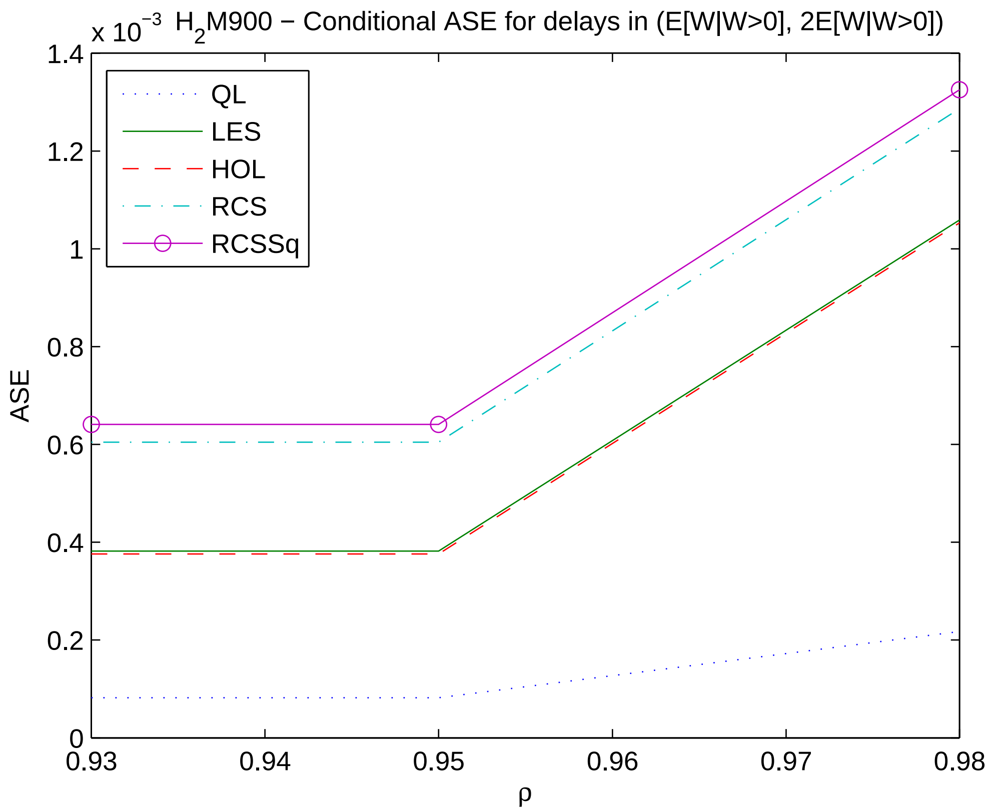

Figure 59: Conditional ASE for the alternative delay estimators in the $H_2/M/900$ model for actual delays in $(E[\widehat{W|W} > 0], 2E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

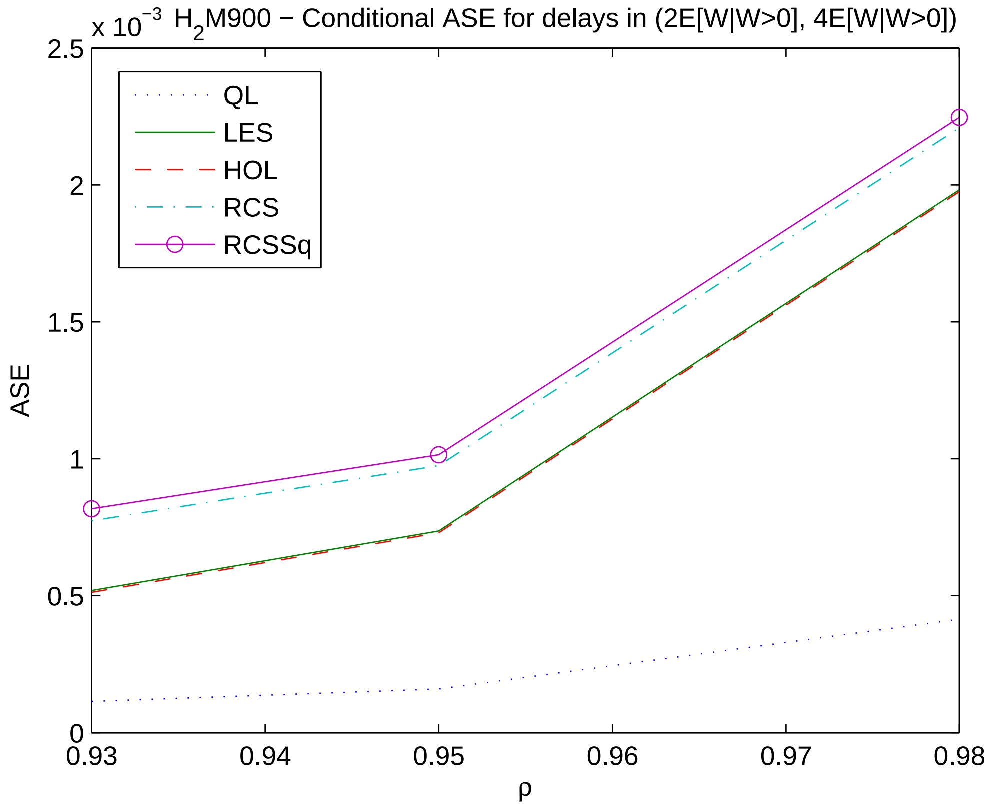

Figure 60: Conditional ASE for the alternative delay estimators in the $H_2/M/900$ model for actual delays in $(2E[\widehat{W|W} > 0], 4E[\widehat{W|W} > 0])$, as a function of the traffic intensity ρ

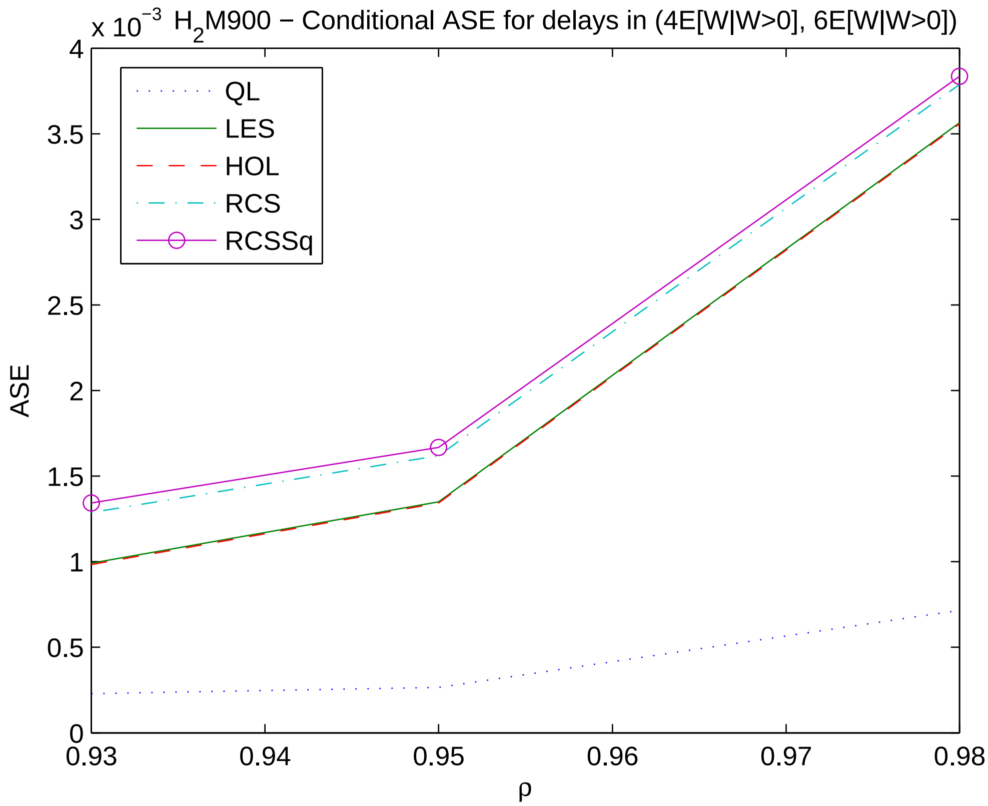

Figure 61: Conditional ASE for the alternative delay estimators in the $H_2/M/900$ model for actual delays in $(4E[\widehat{W|W > 0}], 6E[\widehat{W|W > 0}])$, as a function of the traffic intensity ρ

ASE in the $M/M/s$ model with $s = 1$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.95	42.24 ± 0.766	42	0.571
0.93	30.56 ± 0.371	30.57	-0.0327
0.9	21.76 ± 0.186	22	-1.09
0.85	15.07 ± 0.093	15.33	-1.70

Table 62: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $M/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the M/M/s model with $s = 10$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	1.005 ± 0.0416	1.02	-1.44
0.95	0.4162 ± 0.00399	0.42	-0.905
0.93	0.3033 ± 0.00322	0.3057	-0.785
0.9	0.2185 ± 0.00329	0.22	-0.682
0.85	0.1499 ± 0.000757	0.1533	-2.22

Table 63: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the M/M/s model with s = 100			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	1.023×10^{-2} $\pm 5.04 \times 10^{-4}$	1.02×10^{-2}	0.294
0.95	4.258×10^{-3} $\pm 5.85 \times 10^{-5}$	4.2×10^{-3}	1.38
0.93	3.069×10^{-3} $\pm 2.53 \times 10^{-5}$	3.057×10^{-3}	0.393
0.90	2.168×10^{-3} $\pm 5.87 \times 10^{-5}$	2.2×10^{-3}	-1.45

Table 64: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the M/M/s model with $s = 400$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	6.424×10^{-4} $\pm 2.46 \times 10^{-5}$	6.375×10^{-4}	0.769
0.95	2.646×10^{-4} $\pm 1.11 \times 10^{-5}$	2.625×10^{-4}	0.8
0.93	1.951×10^{-4} $\pm 7.18 \times 10^{-6}$	1.911×10^{4}	2.09
0.9	1.401×10^{-4} $\pm 7.31 \times 10^{-6}$	1.375×10^{-4}	1.89

Table 65: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the M/M/s model with $s = 900$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.99	2.668×10^{-4} $\pm 1.85 \times 10^{-5}$	2.494×10^{-4}	6.98
0.98	1.286×10^{-4} $\pm 4.24 \times 10^{-6}$	1.259×10^{-4}	2.14
0.95	5.219×10^{-5} $\pm 3.85 \times 10^{-6}$	5.185×10^{-5}	0.656
0.93	3.821×10^{-5} $\pm 4.09 \times 10^{-6}$	3.774×10^{-5}	2.06

Table 66: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE *in the* $D/M/s$ *model with* $s = 1$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.95	11.61 ± 0.146	12.11	-4.14
0.93	8.791 ± 0.0780	9.310	-5.57
0.9	6.644 ± 0.0372	7.197	-7.68
0.85	4.955 ± 0.0178	5.551	-10.7

Table 67: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the D/M/s model with $s = 10$

ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	2.633×10^{-1} $\pm 8.34 \times 10^{-3}$	2.684×10^{-1}	-1.92
0.95	1.158×10^{-1} $\pm 1.79 \times 10^{-3}$	1.210×10^{-1}	-4.32
0.93	8.759×10^{-2} $\pm 1.05 \times 10^{-3}$	9.305×10^{-2}	-5.89
0.9	6.630×10^{-2} $\pm 5.70 \times 10^{-4}$	7.197×10^{-2}	-7.88
0.85	4.943×10^{-2} $\pm 2.57 \times 10^{-4}$	5.552×10^{-2}	-11.0

Table 68: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE *in the* $D/M/s$ *model with* $s = 100$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	2.624×10^{-3} $\pm 5.02 \times 10^{-5}$	2.674×10^{-3}	-1.88
0.95	1.154×10^{-3} $\pm 1.78 \times 10^{-5}$	1.203×10^{-3}	-4.12
0.93	8.710×10^{-4} $\pm 1.58 \times 10^{-5}$	9.249×10^{-4}	-5.82
0.9	6.650×10^{-4} $\pm 1.56 \times 10^{-5}$	7.191×10^{-4}	-7.53

Table 69: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE *in the* $D/M/s$ *model with* $s = 400$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.99	3.130×10^{-4} $\pm 1.15 \times 10^{-5}$	3.155×10^{-4}	-0.791
0.98	1.649×10^{-4} $\pm 3.34 \times 10^{-6}$	1.678×10^{-4}	-1.75
0.95	7.248×10^{-5} $\pm 2.41 \times 10^{-6}$	7.491×10^{-5}	-3.24
0.93	5.497×10^{-5} $\pm 2.55 \times 10^{-6}$	5.844×10^{-5}	-5.95
0.9	3.988×10^{-5} $\pm 5.45 \times 10^{-6}$	4.396×10^{-5}	-9.29

Table 70: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the $D/M/s$ model with $s = 900$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.99	6.199×10^{-5} $\pm 2.59 \times 10^{-6}$	6.235×10^{-5}	-0.586
0.98	3.272×10^{-5} $\pm 9.82 \times 10^{-7}$	3.338×10^{-5}	-1.99
0.95	1.531×10^{-5} $\pm 1.47 \times 10^{-6}$	1.582×10^{-5}	-3.24
0.93	1.126×10^{-5} $\pm 1.73 \times 10^{-6}$	1.202×10^{-5}	-6.34

Table 71: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the $H_2/M/s$ model with $s = 1$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.95	226.4 ±5.14	230.3	-1.69
0.93	154.4 ±2.94	158.3	-2.46
0.9	101.3 ±2.32	104.7	-3.25
0.85	59.99 ±0.515	62.81	-4.49

Table 72: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $H_2/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the $H_2/M/s$ model with $s = 10$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	6.259 $\pm$0.404	6.273	-0.223
0.95	2.23 $\pm$0.0466	2.264	-1.50
0.93	1.542 $\pm$0.0346	1.581	-2.47
0.9	1.012 $\pm$0.0181	1.044	-3.07
0.85	0.6004 $\pm$0.0121	0.6266	-4.18

Table 73: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE *in the* $H_2/M/s$ *model with* $s = 100$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	6.040×10^{-2} $\pm 3.21 \times 10^{-3}$	6.283×10^{-2}	-3.87
0.95	2.248×10^{-2} $\pm 4.628 \times 10^{-4}$	2.291×10^{-2}	-1.87
0.93	1.553×10^{-2} $\pm 4.38 \times 10^{-4}$	1.587×10^{-2}	-2.14
0.9	1.020×10^{-2} $\pm 2.11 \times 10^{-4}$	1.054×10^{-2}	-3.22

Table 74: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the $H_2/M/s$ model with $s = 400$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	3.736×10^{-3} $\pm 1.05 \times 10^{-4}$	3.643×10^{-3}	2.55
0.95	1.382×10^{-3} $\pm 4.60 \times 10^{-5}$	1.442×10^{-3}	-4.16
0.93	9.74×10^{-4} $\pm 3.70 \times 10^{-5}$	9.964×10^{-4}	-2.25
0.9	6.294×10^{-4} $\pm 2.56 \times 10^{-5}$	6.517×10^{-4}	-3.42

Table 75: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

ASE in the $H_2/M/s$ model with $s = 900$			
ρ	$ASE(LES)$ (simulation)	approx	RPD (percent)
0.98	7.206×10^{-4} $\pm 3.42 \times 10^{-5}$	7.572×10^{-4}	-4.83
0.95	2.713×10^{-4} $\pm 1.72 \times 10^{-5}$	2.798×10^{-4}	-3.03
0.93	1.935×10^{-4} $\pm 1.29 \times 10^{-5}$	1.959×10^{-4}	-1.23

Table 76: Testing the accuracy of the numerical approximation for the average square error (ASE) of the LES delay estimator in the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. Each simulation estimate is shown with the half width of the 95 percent confidence interval.

Comparison of ASE(QL) and ASE(HOL) in the $M/M/s$ model with $s = 1$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.95	20.09	42.35	2.11	2.11	−0.0941
	±0.417	±0.785			
0.93	14.36	30.72	2.14	2.15	−0.499
	±0.186	±0.385			
0.9	9.989	21.96	2.20	2.22	−0.972
	±0.084	±0.206			
0.85	6.678	15.38	2.30	2.35	−2.00
	±0.043	±0.0951			

Table 77: A comparison of the efficiency of the HOL and QL delay estimators for the $M/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $M/M/s$ model with $s = 10$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	0.4954 ± 0.0232	1.005 ± 0.0413	2.03	2.04	−0.556
0.95	0.1980 ± 0.00251	0.4171 ± 0.00416	2.10	2.11	−0.163
0.93	0.1421 ± 0.00133	0.3046 ± 0.00374	2.14	2.15	−0.299
0.9	0.1003 ± 0.00168	0.2204 ± 0.00416	2.20	2.22	−1.02
0.85	6.613×10^{-2} $\pm 3.24 \times 10^{-4}$	1.528×10^{-1} $\pm 1.24 \times 10^{-4}$	2.31	2.35	−1.63

Table 78: A comparison of the efficiency of the HOL and QL delay estimators for the $M/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $M/M/s$ model with $s = 100$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	5.033×10^{-3}	1.023×10^{-2}	2.03	2.04	-0.363
	2.04 $\pm 2.46 \times 10^{-4}$	$\pm 5.09 \times 10^{-4}$			
0.95	2.041×10^{-3}	4.269×10^{-3}	2.09	2.11	-0.871
	$\pm 4.22 \times 10^{-5}$	$\pm 6.03 \times 10^{-5}$			
0.93	1.442×10^{-3}	3.084×10^{-3}	2.14	2.15	-0.526
	$\pm 1.66 \times 10^{-5}$	$\pm 2.63 \times 10^{-5}$			
0.90	9.940×10^{-4}	2.185×10^{-3}	2.20	2.22	-0.982
	$\pm 2.97 \times 10^{-5}$	$\pm 5.96 \times 10^{-5}$			

Table 79: A comparison of the efficiency of the HOL and QL delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $M/M/s$ model with $s = 400$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	3.145×10^{-4}	6.427×10^{-4}	2.04	2.04	0.175
	$\pm 1.35 \times 10^{-5}$	$\pm 2.50 \times 10^{-5}$			
0.95	1.262×10^{-4}	2.651×10^{-4}	2.10	2.11	−0.444
	$\pm 5.51 \times 10^{-6}$	$\pm 1.13 \times 10^{-5}$			
0.93	9.109×10^{-5}	1.960×10^{-4}	2.15	2.15	0.0799
	$\pm 3.77 \times 10^{-6}$	$\pm 7.35 \times 10^{-6}$			
0.9	6.451×10^{-5}	1.414×10^{-4}	2.19	2.22	−1.27
	$\pm 4.43 \times 10^{-6}$	$\pm 7.77 \times 10^{-6}$			

Table 80: A comparison of the efficiency of the HOL and QL delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $M/M/s$ model with $s = 900$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.99	1.320×10^{-4}	2.669×10^{-4}	2.02	2.02	0.197
	$\pm 8.86 \times 10^{-6}$	$\pm 1.86 \times 10^{-5}$			
0.98	6.334×10^{-5}	1.287×10^{-4}	2.03	2.04	−0.397
	$\pm 2.15 \times 10^{-6}$	$\pm 4.25 \times 10^{-6}$			
0.95	2.484×10^{-5}	5.229×10^{-5}	2.11	2.11	−0.234
	$\pm 1.79 \times 10^{-6}$	$\pm 3.95 \times 10^{-6}$			
0.93	1.752×10^{-5}	3.834×10^{-5}	2.19	2.15	1.78
	$\pm 1.94 \times 10^{-6}$	$\pm 4.11 \times 10^{-6}$			

Table 81: A comparison of the efficiency of the HOL and QL delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $D/M/s$ model with $s = 1$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.95	10.13	11.61	1.15	1.05	9.15
	±0.148	±0.157			
0.93	7.322	8.794	1.20	1.08	11.2
	±0.0805	±0.0862			
0.9	5.192	6.647	1.28	1.11	15.3
	±0.0377	±0.0409			
0.85	3.528	4.954	1.40	1.18	19.0
	±0.0183	±0.0203			

Table 82: A comparison of the efficiency of the HOL and QL delay estimators for the $D/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $D/M/s$ model with $s = 10$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	2.485×10^{-1}	2.633×10^{-1}	1.06	1.02	3.88
	$\pm 8.36 \times 10^{-3}$	$\pm 8.61 \times 10^{-3}$			
0.95	1.010×10^{-1}	1.157×10^{-1}	1.15	1.05	9.10
	$\pm 1.76 \times 10^{-3}$	$\pm 2.03 \times 10^{-3}$			
0.93	7.301×10^{-2}	8.773×10^{-2}	1.20	1.08	10.8
	$\pm 1.04 \times 10^{-3}$	$\pm 1.32 \times 10^{-3}$			
0.9	5.182×10^{-2}	6.632×10^{-2}	1.28	1.11	15.3
	$\pm 5.77 \times 10^{-4}$	$\pm 9.12 \times 10^{-4}$			
0.85	3.519×10^{-2}	4.935×10^{-2}	1.40	1.18	18.8
	$\pm 2.49 \times 10^{-4}$	$\pm 5.73 \times 10^{-4}$			

Table 83: A comparison of the efficiency of the HOL and QL delay estimators for the $D/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $D/M/s$ model with $s = 100$					
ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	2.476×10^{-3}	2.624×10^{-3}	1.06	1.02	3.90
	$\pm 5.04 \times 10^{-5}$	$\pm 5.14 \times 10^{-5}$			
0.95	1.007×10^{-3}	1.153×10^{-3}	1.14	1.05	9.05
	$\pm 1.70 \times 10^{-5}$	$\pm 1.84 \times 10^{-5}$			
0.93	7.250×10^{-4}	8.713×10^{-4}	1.20	1.08	11.3
	$\pm 1.58 \times 10^{-5}$	$\pm 1.70 \times 10^{-5}$			
0.9	5.189×10^{-4}	6.641×10^{-4}	1.28	1.11	15.3
	$\pm 1.52 \times 10^{-5}$	$\pm 1.67 \times 10^{-5}$			

Table 84: A comparison of the efficiency of the HOL and QL delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $D/M/s$ model with $s = 400$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.99	3.035×10^{-4}	3.129×10^{-4}	1.03	1.01	2.08
	$\pm 1.16 \times 10^{-5}$	$\pm 1.17 \times 10^{-5}$			
0.98	1.556×10^{-4}	1.648×10^{-4}	1.06	1.02	3.84
	$\pm 3.37 \times 10^{-6}$	$\pm 3.42 \times 10^{-6}$			
0.95	6.329×10^{-5}	7.245×10^{-5}	1.14	1.05	9.02
	$\pm 2.41 \times 10^{-6}$	$\pm 2.50 \times 10^{-6}$			
0.93	4.583×10^{-5}	5.511×10^{-5}	1.20	1.08	11.3
	$\pm 2.41 \times 10^{-6}$	$\pm 2.64 \times 10^{-6}$			
0.9	3.063×10^{-5}	3.931×10^{-5}	1.28	1.11	15.6
	$\pm 4.88 \times 10^{-6}$	$\pm 5.81 \times 10^{-6}$			

Table 85: A comparison of the efficiency of the HOL and QL delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $D/M/s$ model with $s = 900$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.99	6.013×10^{-5}	6.198×10^{-5}	1.03	1.01	2.06
	$\pm 2.60 \times 10^{-6}$	$\pm 2.62 \times 10^{-6}$			
0.98	3.088×10^{-5}	3.088×10^{-5}	1.00	1.02	-1.96
	$\pm 9.77 \times 10^{-7}$	$\pm 1.00 \times 10^{-7}$			
0.95	1.351×10^{-5}	1.532×10^{-5}	1.13	1.05	8.08
	$\pm 1.50 \times 10^{-6}$	$\pm 1.52 \times 10^{-6}$			
0.93	9.424×10^{-6}	1.125×10^{-5}	1.19	1.08	10.5
	$\pm 1.68 \times 10^{-6}$	$\pm 1.77 \times 10^{-6}$			

Table 86: A comparison of the efficiency of the HOL and QL delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $H_2/M/s$ model with $s = 1$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.95	48.66	226.5	4.65	5.26	−11.5
	±1.13	±5.23			
0.93	34.33	154.4	4.50	5.38	−16.4
	±0.625	±2.94			
0.9	23.48	101.4	4.32	5.56	−22.3
	±0.366	±2.36			
0.85	14.95	60.15	4.02	5.88	−31.6
	±0.104	±0.530			

Table 87: A comparison of the efficiency of the HOL and QL delay estimators for the $H_2/M/s$ queue with $s = 1$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $H_2/M/s$ model with $s = 10$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	1.284 ±0.0690	6.259 ±0.410	4.87	4.59	6.20
0.95	0.4812 ±0.00811	2.231 ±0.0482	4.64	5.26	−11.9
0.93	0.3424 ±0.00691	1.5436 ±0.0370	4.51	5.38	−16.20
0.9	0.2344 ±0.00359	1.013 ±0.0203	4.32	5.56	−22.3
0.85	0.1498 ±0.00217	0.6022 ±0.0133	4.02	5.88	−31.6

Table 88: A comparison of the efficiency of the HOL and QL delay estimators for the $H_2/M/s$ queue with $s = 10$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $H_2/M/s$ model with $s = 100$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	1.238×10^{-2}	6.041×10^{-2}	4.88	5.10	−4.32
	$\pm 6.95 \times 10^{-4}$	$\pm 3.22 \times 10^{-3}$			
0.95	4.815×10^{-3}	2.248×10^{-2}	4.67	5.26	−11.2
	$\pm 9.47 \times 10^{-5}$	$\pm 4.72 \times 10^{-4}$			
0.93	3.444×10^{-3}	1.554×10^{-2}	4.51	5.38	−16.1
	$\pm 9.45 \times 10^{-5}$	$\pm 4.44 \times 10^{-4}$			
0.9	2.350×10^{-3}	1.021×10^{-2}	4.34	5.56	−21.9
	4.03×10^{-5}	2.12×10^{-4}			

Table 89: A comparison of the efficiency of the HOL and QL delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(LES)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(LES) in the $H_2/M/s$ model with $s = 400$					
ρ	ASE(QL)	ASE(LES)	ASE(LES)/ASE(QL) (*sim*)	ASE(LES)/ASE(QL) (*approx*)	RPD
0.98	7.676×10^{-4} $\pm 2.34 \times 10^{-5}$	3.736×10^{-3} $\pm 1.05 \times 10^{-4}$	4.87	5.10	−4.57
0.95	3.018×10^{-4} $\pm 8.57 \times 10^{-6}$	1.382×10^{-3} $\pm 4.67 \times 10^{-5}$	4.58	5.26	−12.9
0.93	2.163×10^{-4} 1.07×10^{-5}	9.752×10^{-4} 3.79×10^{-5}	4.51	5.38	−16.2
0.9	1.443×10^{-4} $\pm 4.09 \times 10^{-6}$	6.300×10^{-4} $\pm 2.59 \times 10^{-5}$	4.37	5.55	−21.3

Table 90: A comparison of the efficiency of the HOL and QL delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

Comparison of ASE(QL) and ASE(HOL) in the $H_2/M/s$ model with $s = 900$

ρ	ASE(QL)	ASE(HOL)	ASE(HOL)/ASE(QL) (*sim*)	ASE(HOL)/ASE(QL) (*approx*)	RPD
0.98	1.487×10^{-4}	7.205×10^{-4}	4.85	5.10	-4.99
	$\pm 7.56 \times 10^{-6}$	$\pm 3.45 \times 10^{-5}$			
0.95	5.826×10^{-5}	2.713×10^{-4}	4.66	5.26	-11.5
	$\pm 3.70 \times 10^{-6}$	$\pm 1.75 \times 10^{-5}$			
0.93	4.292×10^{-5}	1.936×10^{-4}	4.51	5.38	-16.2
	$\pm 2.12 \times 10^{-6}$	$\pm 1.30 \times 10^{-5}$			

Table 91: A comparison of the efficiency of the HOL and QL delay estimators for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included are corresponding values of the approximation ASE(HOL)/ASE(QL) and the relative percent difference RPD.

ASE in the M/M/s model with $s = 100$						
ρ	$ASE(RCS)$	$ASE(RCS - s)$	$ASE(RCS - 4\sqrt{(s)})$	$ASE(RCS - 2\sqrt{(s)})$	$ASE(RCS - \sqrt{(s)})$	$ASE(RCS - log(s))$
0.98	9.439×10^{-3} $\pm 3.13 \times 10^{-4}$	9.439×10^{-3} $\pm 3.13 \times 10^{-4}$	9.439×10^{-3} $\pm 3.13 \times 10^{-4}$	9.452×10^{-3}(**0.138**) $\pm 3.13 \times 10^{-4}$	9.779×10^{-3}(**3.60**) $\pm 3.18 \times 10^{-4}$	1.548×10^{-2}(**64.0**) $\pm 4.50 \times 10^{-4}$
0.97	8.070×10^{-3} $\pm 1.71 \times 10^{-4}$	8.070×10^{-3} $\pm 1.71 \times 10^{-4}$	8.070×10^{-3} $\pm 1.71 \times 10^{-4}$	8.083×10^{-3}(**0.161**) $\pm 1.72 \times 10^{-4}$	8.395×10^{-3}(**4.03**) $\pm 1.80 \times 10^{-4}$	1.388×10^{-2}(**72.0**) $\pm 3.34 \times 10^{-4}$
0.95	6.280×10^{-3} $\pm 1.78 \times 10^{-4}$	6.280×10^{-3} $\pm 1.78 \times 10^{-4}$	6.280×10^{-3} $\pm 1.78 \times 10^{-4}$	6.295×10^{-3}(**0.239**) $\pm 1.79 \times 10^{-4}$	6.571×10^{-3}(**4.63**) $\pm 1.82 \times 10^{-4}$	1.135×10^{-2}(**80.7**) $\pm 3.01 \times 10^{-4}$
0.93	4.908×10^{-3} $\pm 1.22 \times 10^{-4}$	4.908×10^{-3} $\pm 1.22 \times 10^{-4}$	4.908×10^{-3} $\pm 1.22 \times 10^{-4}$	4.918×10^{-3}(**0.204**) $\pm 1.23 \times 10^{-4}$	5.161×10^{-3}(**5.15**) $\pm 1.23 \times 10^{-4}$	9.017×10^{-3}(**83.7**) $\pm 1.91 \times 10^{-4}$
0.9	3.897×10^{-3} $\pm 9.62 \times 10^{-5}$	3.897×10^{-3} $\pm 9.62 \times 10^{-5}$	3.897×10^{-3} $\pm 9.62 \times 10^{-5}$	3.906×10^{-3}(**0.696**) $\pm 9.62 \times 10^{-5}$	4.108×10^{-3}(**5.41**) $\pm 1.01 \times 10^{-4}$	6.892×10^{-3}(**76.9**) $\pm 1.92 \times 10^{-4}$

Table 92: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $M/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the M/M/s model with $s = 400$						
ρ	$ASE(RCS)$	$ASE(RCS-s)$	$ASE(RCS-4\sqrt(s))$	$ASE(RCS-2\sqrt(s))$	$ASE(RCS-\sqrt(s))$	$ASE(RCS-log(s))$
0.98	7.298×10^{-4}	7.298×10^{-4}	7.298×10^{-4}	7.318×10^{-4}(**0.274**)	7.696×10^{-4}(**5.45**)	1.663×10^{-3}(**128**)
	$\pm 3.45 \times 10^{-5}$	$\pm 3.45 \times 10^{-5}$	$\pm 3.45 \times 10^{-5}$	$\pm 3.45 \times 10^{-5}$	$\pm 3.59 \times 10^{-5}$	$\pm 7.14 \times 10^{-5}$
0.97	6.457×10^{-4}	6.457×10^{-4}	6.457×10^{-4}	6.475×10^{-4}(**0.279**)	6.849×10^{-4}(**6.07**)	1.512×10^{-3}(**134**)
	$\pm 1.37 \times 10^{-5}$	$\pm 1.37 \times 10^{-5}$	$\pm 1.37 \times 10^{-5}$	$\pm 1.38 \times 10^{-5}$	$\pm 1.37 \times 10^{-5}$	$\pm 3.69 \times 10^{-5}$
0.95	5.087×10^{-4}	5.087×10^{-4}	5.087×10^{-4}	5.102×10^{-4}(**0.295**)	5.416×10^{-4}(**6.47**)	1.184×10^{-3}(**133**)
	$\pm 2.61 \times 10^{-5}$	$\pm 2.61 \times 10^{-5}$	$\pm 2.61 \times 10^{-5}$	$\pm 2.60 \times 10^{-5}$	$\pm 2.72 \times 10^{-5}$	$\pm 7.21 \times 10^{-5}$
0.93	4.118×10^{-4}	4.118×10^{-4}	4.118×10^{-4}	4.124×10^{-4}(**0.146**)	4.351×10^{-4}(**5.66**)	9.242×10^{-4}(**124**)
	$\pm 2.07 \times 10^{-5}$	$\pm 2.07 \times 10^{-5}$	$\pm 2.07 \times 10^{-5}$	$\pm 2.08 \times 10^{-5}$	$\pm 1.98 \times 10^{-5}$	$\pm 6.68 \times 10^{-5}$

Table 93: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $M/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the M/M/s model with $s = 900$						
ρ	$ASE(RCS)$	$ASE(RCS-s)$	$ASE(RCS-4\sqrt{(s)})$	$ASE(RCS-2\sqrt{(s)})$	$ASE(RCS-\sqrt{(s)})$	$ASE(RCS-log(s))$
0.98	1.638×10^{-4}	1.638×10^{-4}	1.638×10^{-4}	1.642×10^{-4}(**0.244**)	1.753×10^{-4}(**7.02**)	5.262×10^{-4}(**221**)
	$\pm 4.66 \times 10^{-6}$	$\pm 4.66 \times 10^{-6}$	$\pm 4.66 \times 10^{-6}$	$\pm 4.73 \times 10^{-6}$	$\pm 5.11 \times 10^{-6}$	$\pm 1.24 \times 10^{-5}$
0.97	1.474×10^{-4}	1.474×10^{-4}	1.474×10^{-4}	1.479×10^{-4}(**0.339**)	1.586×10^{-4}(**7.60**)	4.717×10^{-4}(**220**)
	$\pm 6.21 \times 10^{-6}$	$\pm 6.21 \times 10^{-6}$	$\pm 6.21 \times 10^{-6}$	$\pm 6.19 \times 10^{-6}$	$\pm 6.45 \times 10^{-6}$	$\pm 3.46 \times 10^{-5}$
0.95	1.185×10^{-4}	1.185×10^{-4}	1.185×10^{-4}	1.190×10^{-4}(**0.421**)	1.272×10^{-4}(**7.34**)	3.477×10^{-4}(**193**)
	$\pm 8.44 \times 10^{-6}$	$\pm 8.44 \times 10^{-6}$	$\pm 8.44 \times 10^{-6}$	$\pm 8.45 \times 10^{-6}$	$\pm 8.71 \times 10^{-6}$	$\pm 4.11 \times 10^{-5}$
0.93	1.079×10^{-4}	1.079×10^{-4}	1.079×10^{-4}	1.084×10^{-4}(**0.463**)	1.149×10^{-4}(**6.49**)	2.847×10^{-4}(**164**)
	$\pm 1.91 \times 10^{-5}$	$\pm 1.91 \times 10^{-5}$	$\pm 1.91 \times 10^{-5}$	$\pm 1.92 \times 10^{-5}$	$\pm 2.11 \times 10^{-5}$	$\pm 8.02 \times 10^{-5}$

Table 94: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $M/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE *in the* $H_2/M/s$ *model with* $s = 100$						
ρ	$ASE(RCS)$	$ASE(RCS - s)$	$ASE(RCS - 4\sqrt(s))$	$ASE(RCS - 2\sqrt(s))$	$ASE(RCS - \sqrt(s))$	$ASE(RCS - log(s))$
0.98	2.439×10^{-2} $\pm 4.84 \times 10^{-4}$	2.439×10^{-2} $\pm 4.84 \times 10^{-4}$	2.439×10^{-2} $\pm 4.84 \times 10^{-4}$	2.442×10^{-2}(**0.123**) $\pm 4.88 \times 10^{-4}$	2.511×10^{-2}(**2.95**) $\pm 4.92 \times 10^{-4}$	3.724×10^{-2}(**52.7**) $\pm 6.81 \times 10^{-4}$
0.97	2.229×10^{-2} $\pm 4.70 \times 10^{-4}$	2.229×10^{-2} $\pm 4.70 \times 10^{-4}$	2.229×10^{-2} $\pm 4.70 \times 10^{-4}$	2.229×10^{-2}(**0.141**) $\pm 4.73 \times 10^{-4}$	2.367×10^{-2}(**3.28**) $\pm 4.73 \times 10^{-4}$	3.566×10^{-2}(**55.6**) $\pm 5.80 \times 10^{-4}$
0.95	1.989×10^{-2} $\pm 3.67 \times 10^{-4}$	1.989×10^{-2} $\pm 3.67 \times 10^{-4}$	1.989×10^{-2} $\pm 3.67 \times 10^{-4}$	1.992×10^{-2}(**0.136**) $\pm 3.67 \times 10^{-4}$	2.058×10^{-2}(**3.48**) $\pm 3.64 \times 10^{-4}$	3.175×10^{-2}(**59.6**) $\pm 5.45 \times 10^{-4}$
0.93	1.715×10^{-2} $\pm 3.56 \times 10^{-4}$	1.715×10^{-2} $\pm 3.56 \times 10^{-4}$	1.715×10^{-2} $\pm 3.56 \times 10^{-4}$	1.718×10^{-2}(**0.150**) $\pm 3.54 \times 10^{-4}$	1.780×10^{-2}(**3.78**) $\pm 3.60 \times 10^{-4}$	2.800×10^{-2}(**63.2**) $\pm 5.89 \times 10^{-4}$
0.90	1.344×10^{-2} $\pm 4.90 \times 10^{-4}$	1.344×10^{-2} $\pm 4.90 \times 10^{-4}$	1.344×10^{-2} $\pm 4.90 \times 10^{-4}$	1.347×10^{-2}(**0.182**) $\pm 4.89 \times 10^{-4}$	1.399×10^{-2}(**4.06**) $\pm 4.99 \times 10^{-4}$	2.233×10^{-2}(**66.3**) $\pm 8.61 \times 10^{-4}$

Table 95: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $H_2/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the $H_2/M/s$ *model with* $s = 400$						
ρ	$ASE(RCS)$	$ASE(RCS-s)$	$ASE(RCS-4\sqrt(s))$	$ASE(RCS-2\sqrt(s))$	$ASE(RCS-\sqrt(s))$	$ASE(RCS-log(s))$
0.98	1.795×10^{-3} $\pm 3.97 \times 10^{-5}$	1.795×10^{-3} $\pm 3.97 \times 10^{-5}$	1.795×10^{-3} $\pm 3.97 \times 10^{-5}$	1.799×10^{-3}(**0.249**) $\pm 3.96 \times 10^{-5}$	1.883×10^{-3}(**4.90**) $\pm 4.13 \times 10^{-5}$	3.634×10^{-3}(**102**) $\pm 7.32 \times 10^{-5}$
0.97	1.733×10^{-3} $\pm 4.06 \times 10^{-5}$	1.733×10^{-3} $\pm 4.06 \times 10^{-5}$	1.733×10^{-3} $\pm 4.06 \times 10^{-5}$	1.737×10^{-3}(**0.242**) $\pm 4.11 \times 10^{-5}$	1.819×10^{-3}(**5.02**) $\pm 4.50 \times 10^{-5}$	3.561×10^{-3}(**106**) $\pm 9.86 \times 10^{-5}$
0.95	1.506×10^{-3} $\pm 4.57 \times 10^{-5}$	1.506×10^{-3} $\pm 4.57 \times 10^{-5}$	1.506×10^{-3} $\pm 4.57 \times 10^{-5}$	1.509×10^{-3}(**0.194**) $\pm 4.57 \times 10^{-5}$	1.579×10^{-3}(**4.86**) $\pm 4.70 \times 10^{-5}$	3.133×10^{-3}(**108**) $\pm 1.10 \times 10^{-4}$
0.93	1.361×10^{-3} $\pm 2.70 \times 10^{-5}$	1.361×10^{-3} $\pm 2.70 \times 10^{-5}$	1.361×10^{-3} $\pm 2.70 \times 10^{-5}$	1.365×10^{-3}(**0.248**) $\pm 2.73 \times 10^{-5}$	1.434×10^{-3}(**5.32**) $\pm 2.94 \times 10^{-5}$	2.848×10^{-3}(**109**) $\pm 6.63 \times 10^{-5}$

Table 96: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $H_2/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the $H_2/M/s$ model with $s = 900$						
ρ	$ASE(RCS)$	$ASE(RCS-s)$	$ASE(RCS-4\sqrt(s))$	$ASE(RCS-2\sqrt(s))$	$ASE(RCS-\sqrt(s))$	$ASE(RCS-log(s))$
0.98	4.216×10^{-4}	4.216×10^{-4}	4.216×10^{-4}	$4.139 \times 10^{-4}(\mathbf{0.307})$	$4.356 \times 10^{-4}(\mathbf{5.56})$	$1.075 \times 10^{-3}(\mathbf{161})$
	$\pm 9.97 \times 10^{-6}$	$\pm 9.97 \times 10^{-6}$	$\pm 9.97 \times 10^{-6}$	$\pm 9.92 \times 10^{-6}$	$\pm 1.03 \times 10^{-5}$	$\pm 2.82 \times 10^{-5}$
0.97	3.914×10^{-4}	3.914×10^{-4}	3.914×10^{-4}	$3.925 \times 10^{-4}(\mathbf{0.300})$	$4.124 \times 10^{-4}(\mathbf{5.37})$	$1.019 \times 10^{-3}(\mathbf{160})$
	$\pm 6.87 \times 10^{-6}$	$\pm 6.87 \times 10^{-6}$	$\pm 6.87 \times 10^{-6}$	$\pm 6.81 \times 10^{-6}$	$\pm 6.96 \times 10^{-6}$	$\pm 1.94 \times 10^{-5}$
0.95	3.543×10^{-4}	3.543×10^{-4}	3.543×10^{-4}	$3.553 \times 10^{-4}(\mathbf{0.304})$	$3.768 \times 10^{-4}(\mathbf{6.37})$	$9.449 \times 10^{-4}(\mathbf{167})$
	$\pm 1.18 \times 10^{-5}$	$\pm 1.18 \times 10^{-5}$	$\pm 1.18 \times 10^{-5}$	$\pm 1.19 \times 10^{-5}$	$\pm 1.23 \times 10^{-5}$	$\pm 3.48 \times 10^{-5}$
0.93	3.213×10^{-4}	3.213×10^{-4}	3.213×10^{-4}	$3.221 \times 10^{-4}(\mathbf{0.243})$	$3.412 \times 10^{-4}(\mathbf{6.16})$	$8.371 \times 10^{-4}(\mathbf{161})$
	$\pm 1.53 \times 10^{-5}$	$\pm 1.53 \times 10^{-5}$	$\pm 1.53 \times 10^{-5}$	$\pm 1.53 \times 10^{-5}$	$\pm 1.68 \times 10^{-5}$	$\pm 5.42 \times 10^{-5}$

Table 97: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $H_2/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the $D/M/s$ model with $s = 100$						
ρ	$ASE(RCS)$	$ASE(RCS-s)$	$ASE(RCS-4\sqrt(s))$	$ASE(RCS-2\sqrt(s))$	$ASE(RCS-\sqrt(s))$	$ASE(RCS-log(s))$
0.98	3.617×10^{-3} $\pm 9.75 \times 10^{-5}$	3.617×10^{-3} $\pm 9.75 \times 10^{-5}$	3.617×10^{-3} $\pm 9.75 \times 10^{-5}$	$3.624 \times 10^{-3}(\mathbf{0.194})$ $\pm 9.75 \times 10^{-5}$	$3.789 \times 10^{-3}(\mathbf{4.76})$ $\pm 9.10 \times 10^{-5}$	$6.678 \times 10^{-3}(\mathbf{84.6})$ $\pm 1.27 \times 10^{-4}$
0.97	2.906×10^{-3} $\pm 6.80 \times 10^{-5}$	2.906×10^{-3} $\pm 6.80 \times 10^{-5}$	2.906×10^{-3} $\pm 6.80 \times 10^{-5}$	$2.913 \times 10^{-3}(\mathbf{0.241})$ $\pm 7.13 \times 10^{-5}$	$3.066 \times 10^{-3}(\mathbf{5.51})$ $\pm 6.81 \times 10^{-5}$	$5.665 \times 10^{-3}(\mathbf{94.9})$ $\pm 1.22 \times 10^{-4}$
0.95	2.200×10^{-3} $\pm 4.40 \times 10^{-5}$	2.200×10^{-3} $\pm 4.40 \times 10^{-5}$	2.200×10^{-3} $\pm 4.40 \times 10^{-5}$	$2.205 \times 10^{-3}(\mathbf{0.227})$ $\pm 4.42 \times 10^{-5}$	$2.337 \times 10^{-3}(\mathbf{6.23})$ $\pm 4.38 \times 10^{-5}$	$4.406 \times 10^{-3}(\mathbf{100})$ $\pm 1.13 \times 10^{-4}$
0.93	1.852×10^{-3} $\pm 4.14 \times 10^{-5}$	1.852×10^{-3} $\pm 4.14 \times 10^{-5}$	1.852×10^{-3} $\pm 4.14 \times 10^{-5}$	$1.856 \times 10^{-3}(\mathbf{0.216})$ $\pm 4.22 \times 10^{-5}$	$1.966 \times 10^{-3}(\mathbf{6.16})$ $\pm 4.24 \times 10^{-5}$	$3.594 \times 10^{-3}(\mathbf{94.1})$ $\pm 1.06 \times 10^{-4}$

Table 98: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $D/M/s$ queue with $s = 100$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE in the $D/M/s$ model with $s = 400$						
ρ	$ASE(RCS)$	$ASE(RCS - s)$	$ASE(RCS - 4\sqrt(s))$	$ASE(RCS - 2\sqrt(s))$	$ASE(RCS - \sqrt(s))$	$ASE(RCS - log(s))$
0.98	3.031×10^{-4}	3.031×10^{-4}	3.031×10^{-4}	3.041×10^{-4}(**0.330**)	3.241×10^{-4}(**6.93**)	7.908×10^{-4}(**161**)
	$\pm 1.32 \times 10^{-5}$	$\pm 1.32 \times 10^{-5}$	$\pm 1.32 \times 10^{-5}$	$\pm 1.38 \times 10^{-5}$	$\pm 1.37 \times 10^{-5}$	$\pm 3.10 \times 10^{-5}$
0.97	2.495×10^{-4}	2.495×10^{-4}	2.495×10^{-4}	2.504×10^{-4}(**0.361**)	2.687×10^{-4}(**7.70**)	6.604×10^{-4}(**165**)
	$\pm 9.33 \times 10^{-6}$	$\pm 9.33 \times 10^{-6}$	$\pm 9.33 \times 10^{-6}$	$\pm 9.75 \times 10^{-6}$	$\pm 1.03 \times 10^{-5}$	$\pm 2.99 \times 10^{-5}$
0.95	1.989×10^{-4}	1.989×10^{-4}	1.989×10^{-4}	1.996×10^{-4}(**0.352**)	2.144×10^{-4}(**7.79**)	4.954×10^{-4}(**149**)
	$\pm 1.27 \times 10^{-5}$	$\pm 1.27 \times 10^{-5}$	$\pm 1.27 \times 10^{-5}$	$\pm 1.33 \times 10^{-5}$	$\pm 1.38 \times 10^{-5}$	$\pm 4.20 \times 10^{-5}$
0.93	1.662×10^{-4}	1.662×10^{-4}	1.662×10^{-4}	1.667×10^{-4}(**0.300**)	1.789×10^{-4}(**7.64**)	3.692×10^{-4}(**122**)
	$\pm 1.88 \times 10^{-5}$	$\pm 1.88 \times 10^{-5}$	$\pm 1.88 \times 10^{-5}$	$\pm 2.02 \times 10^{-5}$	$\pm 2.14 \times 10^{-5}$	$\pm 6.34 \times 10^{-5}$

Table 99: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $D/M/s$ queue with $s = 400$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

ASE *in the* $D/M/s$ *model with* $s = 900$						
ρ	$ASE(RCS)$	$ASE(RCS - s)$	$ASE(RCS - 4\sqrt(s))$	$ASE(RCS - 2\sqrt(s))$	$ASE(RCS - \sqrt(s))$	$ASE(RCS - log(s))$
0.98	7.377×10^{-5}	7.377×10^{-5}	7.377×10^{-5}	$7.411 \times 10^{-5}(\mathbf{0.461})$	$7.987 \times 10^{-5}(\mathbf{8.27})$	$2.744 \times 10^{-4}(\mathbf{272})$
	$\pm 4.01 \times 10^{-6}$	$\pm 4.01 \times 10^{-6}$	$\pm 4.01 \times 10^{-6}$	$\pm 3.87 \times 10^{-6}$	$\pm 4.13 \times 10^{-6}$	$\pm 1.85 \times 10^{-5}$
0.97	6.307×10^{-5}	6.307×10^{-5}	6.307×10^{-5}	$6.339 \times 10^{-5}(\mathbf{0.507})$	$6.878 \times 10^{-5}(\mathbf{9.05})$	$2.309 \times 10^{-4}(\mathbf{266})$
	$\pm 3.83 \times 10^{-6}$	$\pm 3.83 \times 10^{-6}$	$\pm 3.83 \times 10^{-6}$	$\pm 3.89 \times 10^{-6}$	$\pm 4.22 \times 10^{-6}$	$\pm 2.66 \times 10^{-5}$
0.95	5.479×10^{-5}	5.479×10^{-5}	5.479×10^{-5}	$5.510 \times 10^{-5}(\mathbf{0.566})$	$5.955 \times 10^{-5}(\mathbf{8.69})$	$1.792 \times 10^{-4}(\mathbf{227})$
	$\pm 4.83 \times 10^{-6}$	$\pm 4.83 \times 10^{-6}$	$\pm 4.83 \times 10^{-6}$	$\pm 4.88 \times 10^{-6}$	$\pm 5.36 \times 10^{-6}$	$\pm 4.18 \times 10^{-5}$
0.93	3.802×10^{-5}	3.802×10^{-5}	3.802×10^{-5}	$3.833 \times 10^{-5}(\mathbf{0.815})$	$4.117 \times 10^{-5}(\mathbf{8.29})$	$1.008 \times 10^{-4}(\mathbf{165})$
	$\pm 1.03 \times 10^{-5}$	$\pm 1.03 \times 10^{-5}$	$\pm 1.03 \times 10^{-5}$	$\pm 1.03 \times 10^{-5}$	$\pm 1.15 \times 10^{-5}$	$\pm 5.01 \times 10^{-5}$

Table 100: A comparison of the efficiency of the candidate RCS-$f(s)$ delay estimators for the $D/M/s$ queue with $s = 900$ and $\mu = 1$ as a function of the traffic intensity ρ. We report point estimates for the average squared error – (ASE). Each estimate is shown with the half width of the 95 percent confidence interval. Also included in parentheses are the values of the relative percent difference – (RPD).

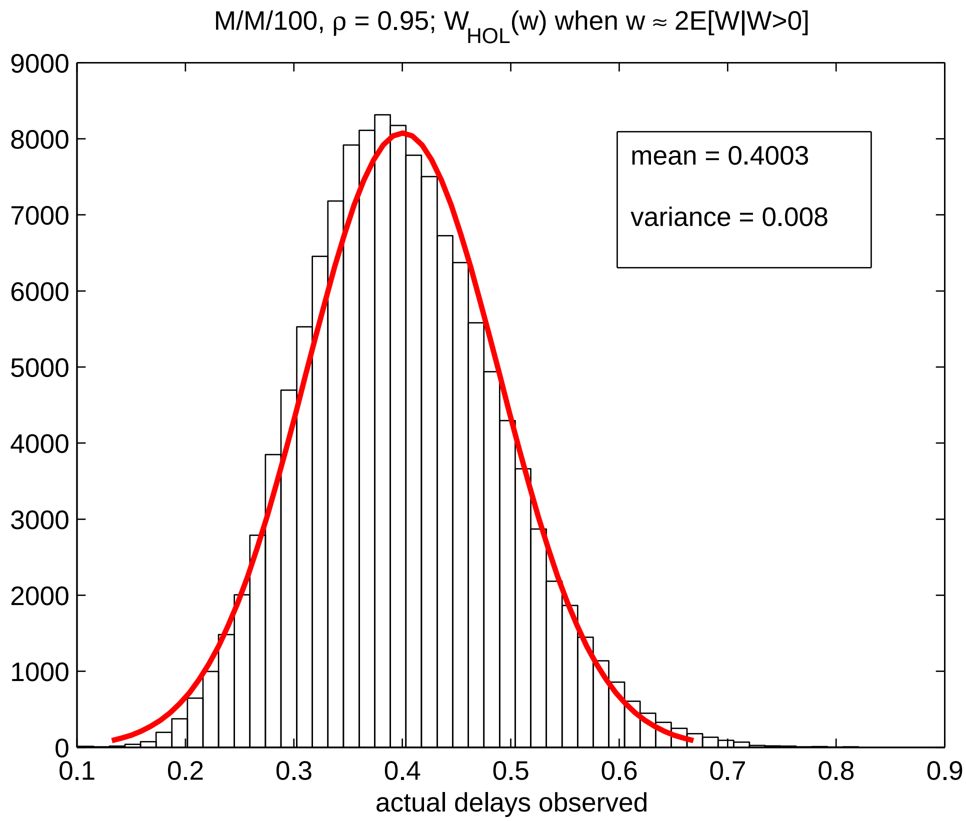

Figure 62: Distribution of $W_{HOL}(w)$ when w $\approx 2E[W|W>0]$ in the $M/M/100$ model when $\rho = 0.95$

{MM100_actual

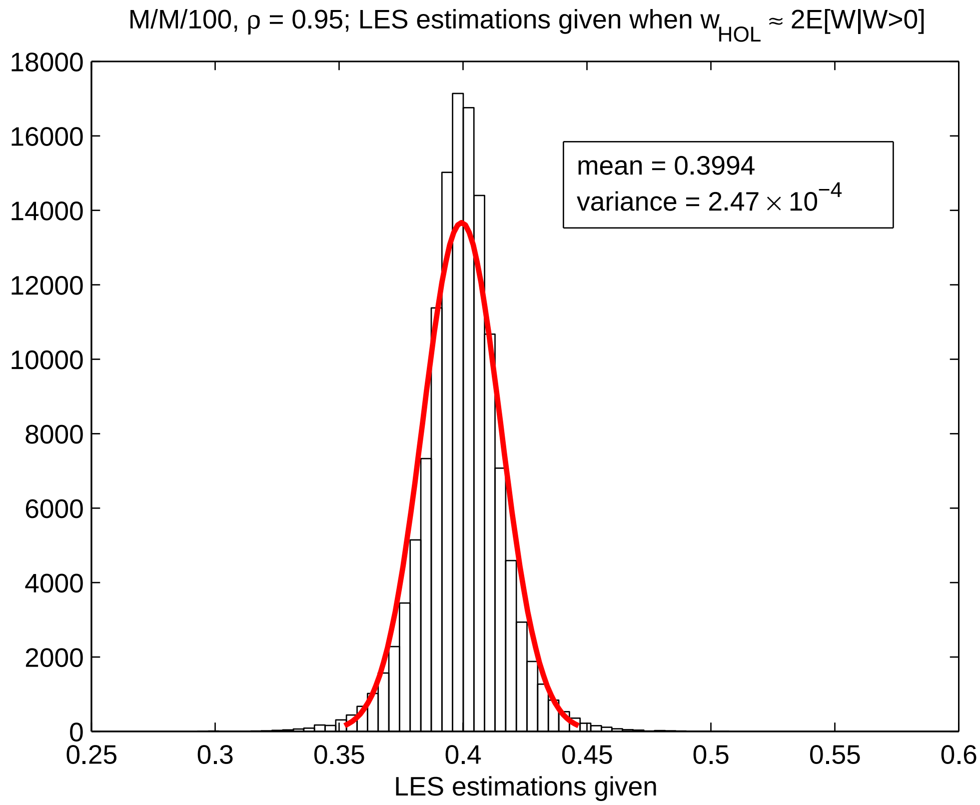

Figure 63: Distribution of the LES delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W>0]$ in the $M/M/100$ model when $\rho = 0.95$

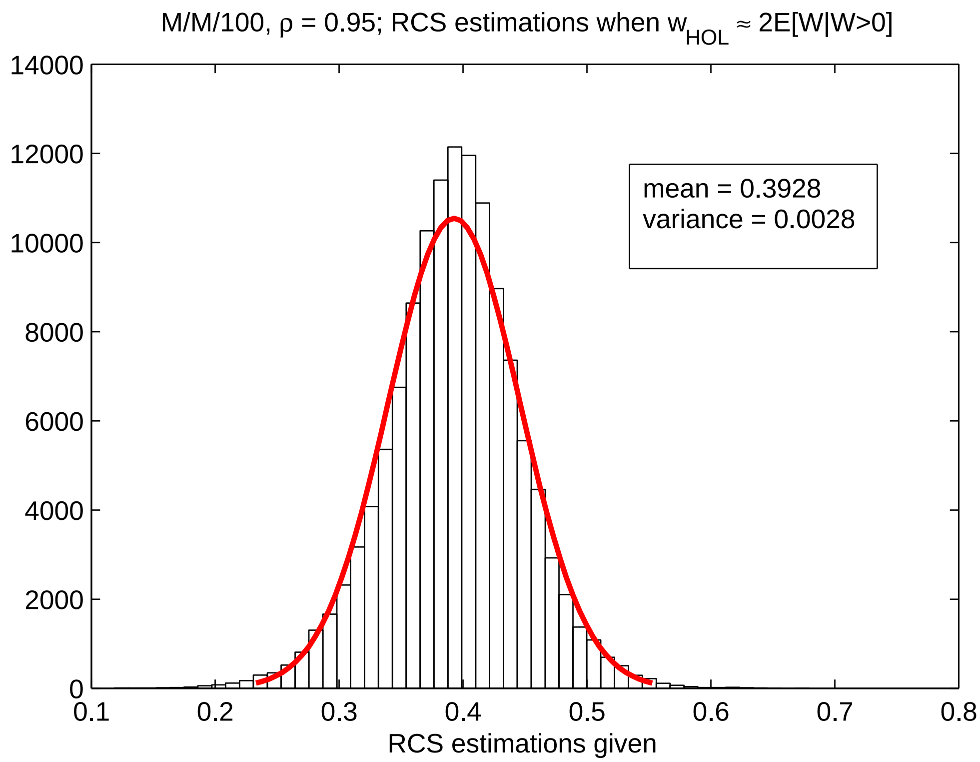

Figure 64: Distribution of the RCS delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W>0]$ in the $M/M/100$ model when $\rho = 0.95$

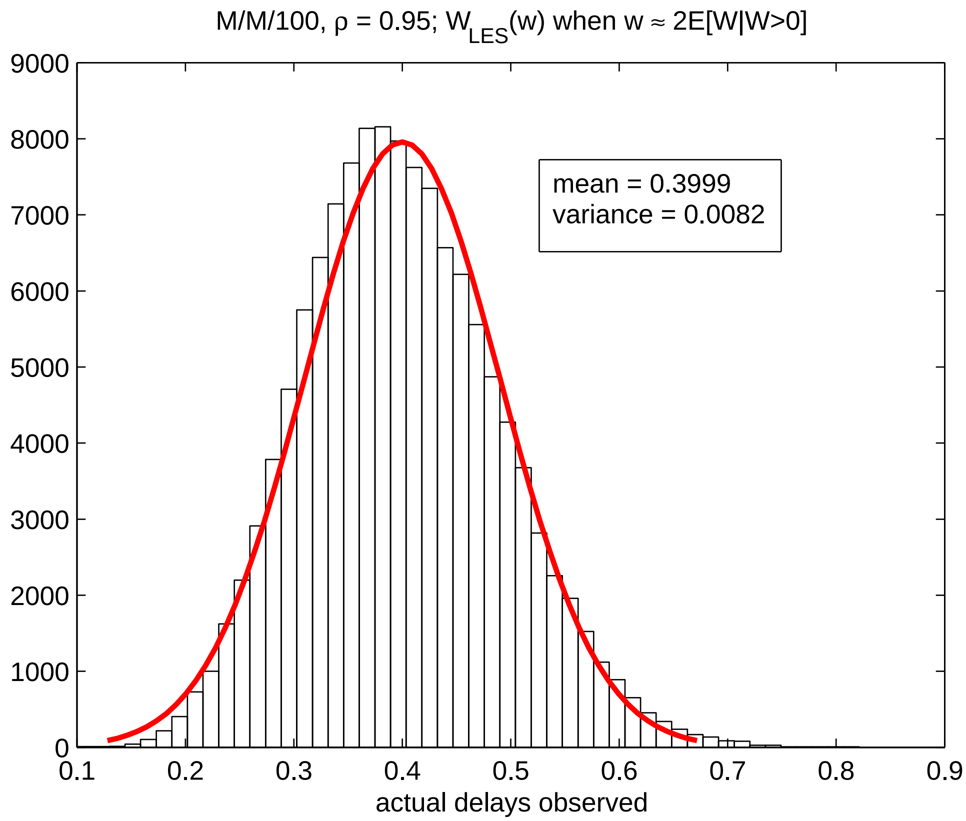

Figure 65: Distribution of the actual delays observed when the LES delay estimation, w, is such that: w $\approx 2E[W|W>0]$ in the $M/M/100$ model when $\rho = 0.95$

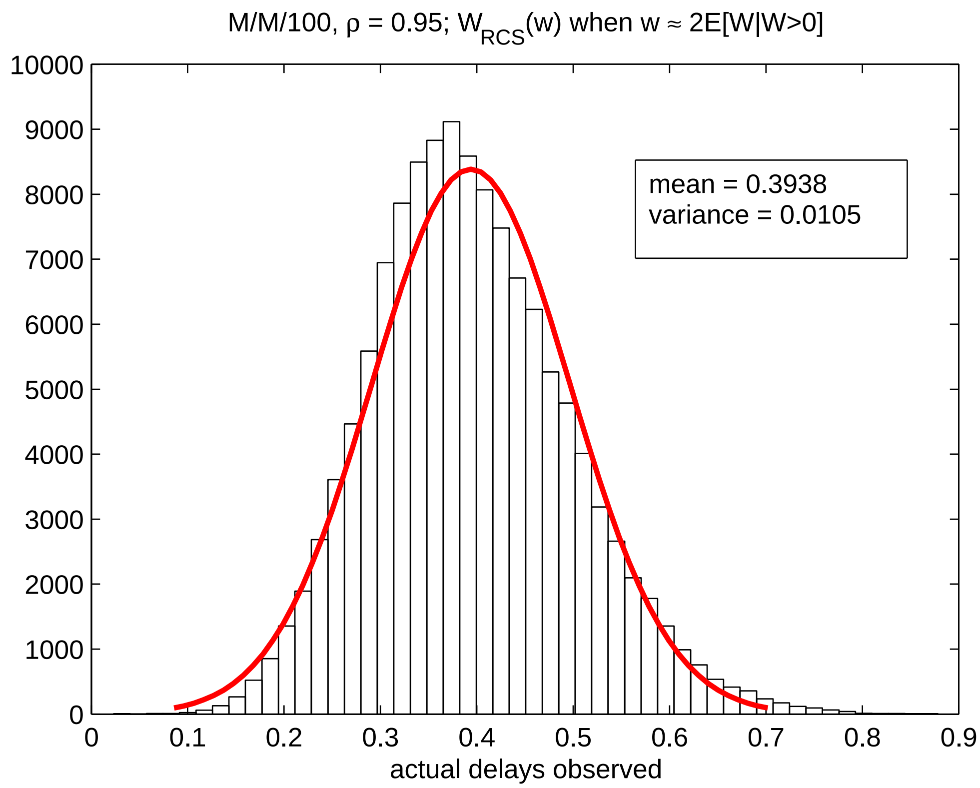

Figure 66: Distribution of the actual delays observed when the RCS delay estimation, w, is such that: w $\approx 2E[W|W>0]$ in the $M/M/100$ model when $\rho = 0.95$

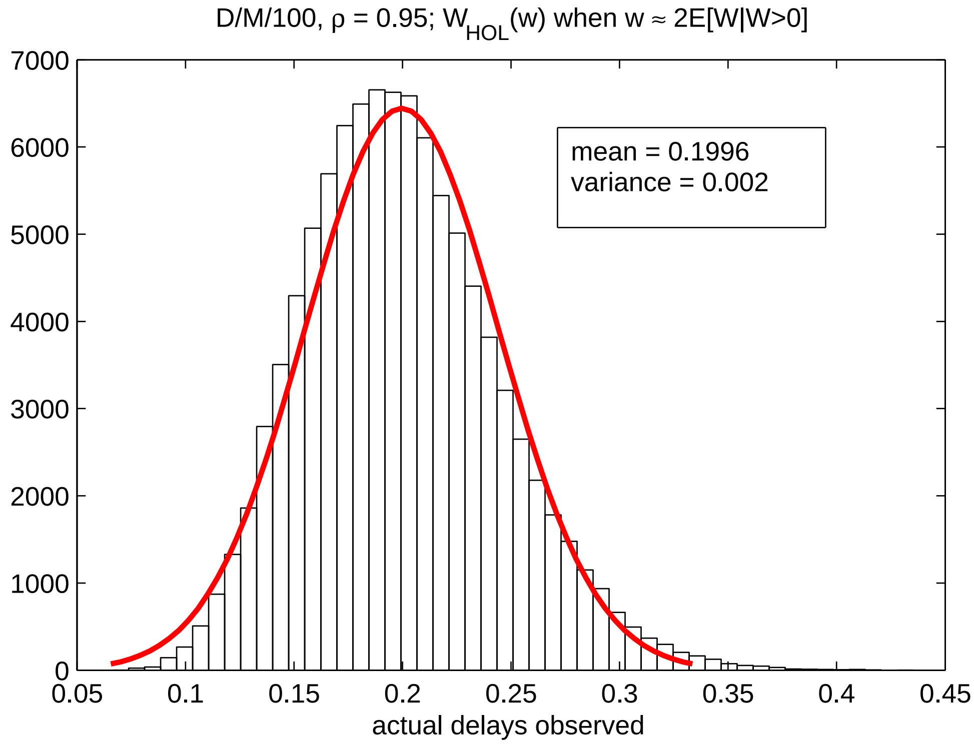

Figure 67: Distribution of $W_{HOL}(w)$ when w $\approx 2E[W|W > 0]$ in the $D/M/100$ model when $\rho = 0.95$

{DM100_actual

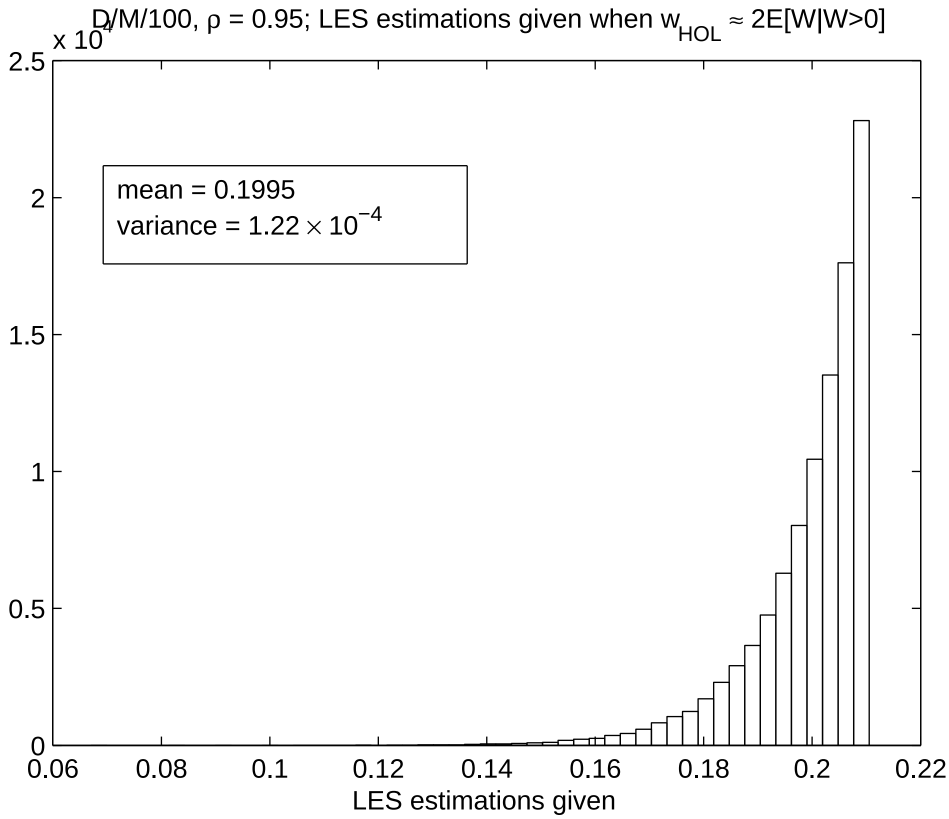

Figure 68: Distribution of the LES delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W > 0]$ in the $D/M/100$ model when $\rho = 0.95$

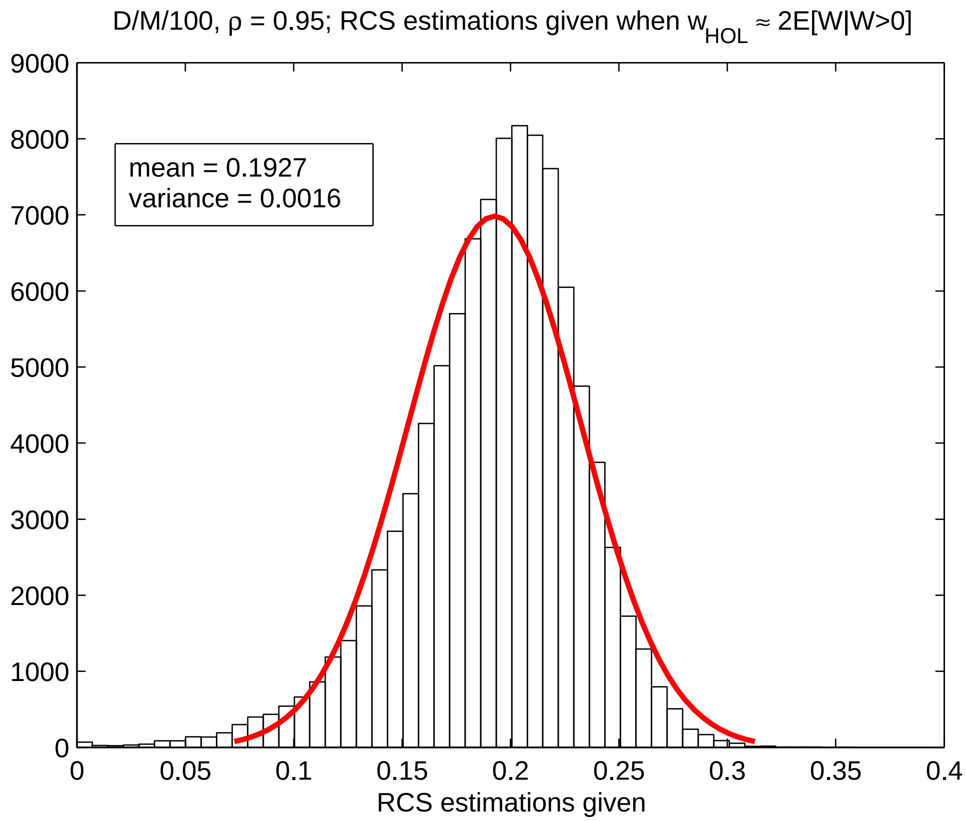

Figure 69: Distribution of the RCS delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W > 0]$ in the $D/M/100$ model when $\rho = 0.95$

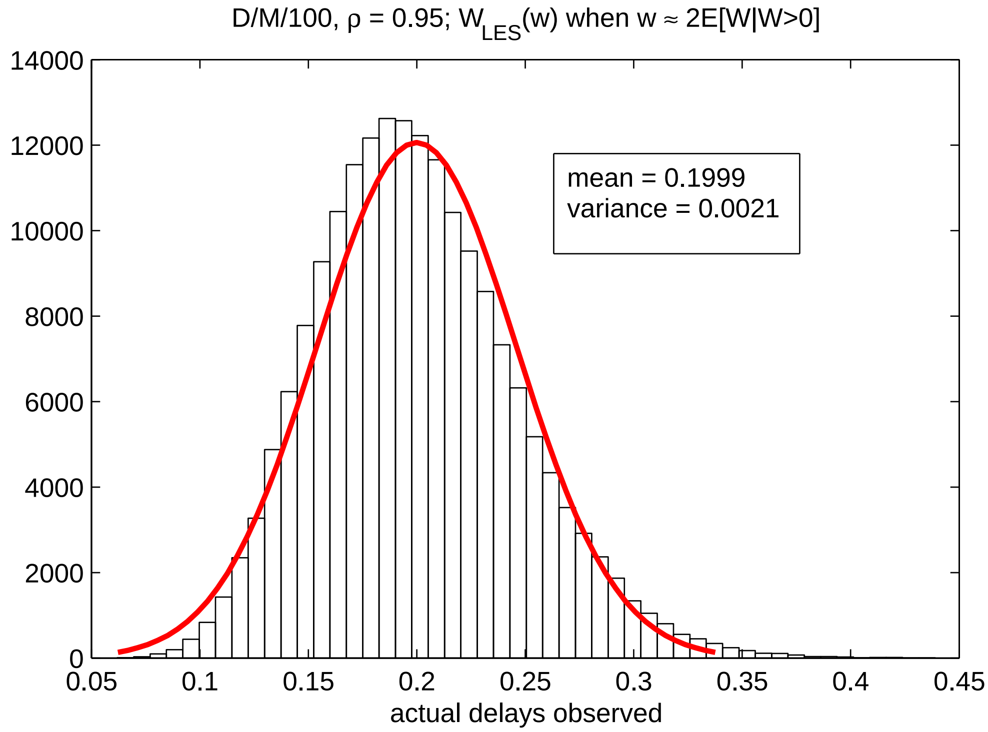

Figure 70: Distribution of the actual delays observed when the LES delay estimation, w, is such that: w $\approx 2E[W|W>0]$ in the $D/M/100$ model when $\rho = 0.95$

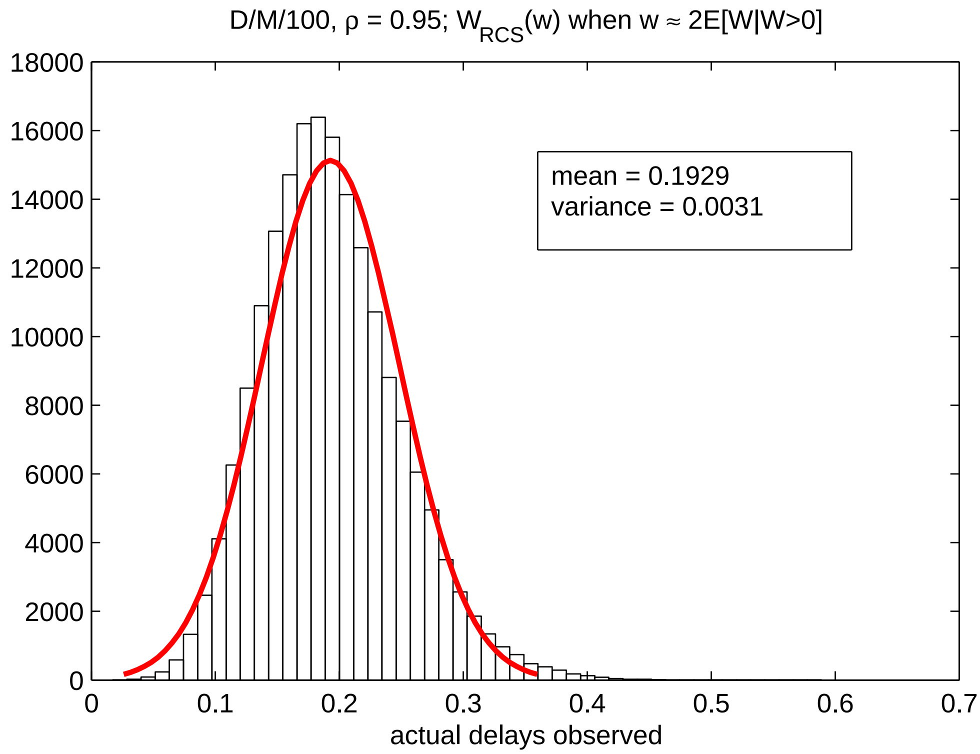

Figure 71: Distribution of the actual delays observed when the RCS delay estimation, w, is such that: w $\approx 2E[W|W > 0]$ in the $D/M/100$ model when $\rho = 0.95$

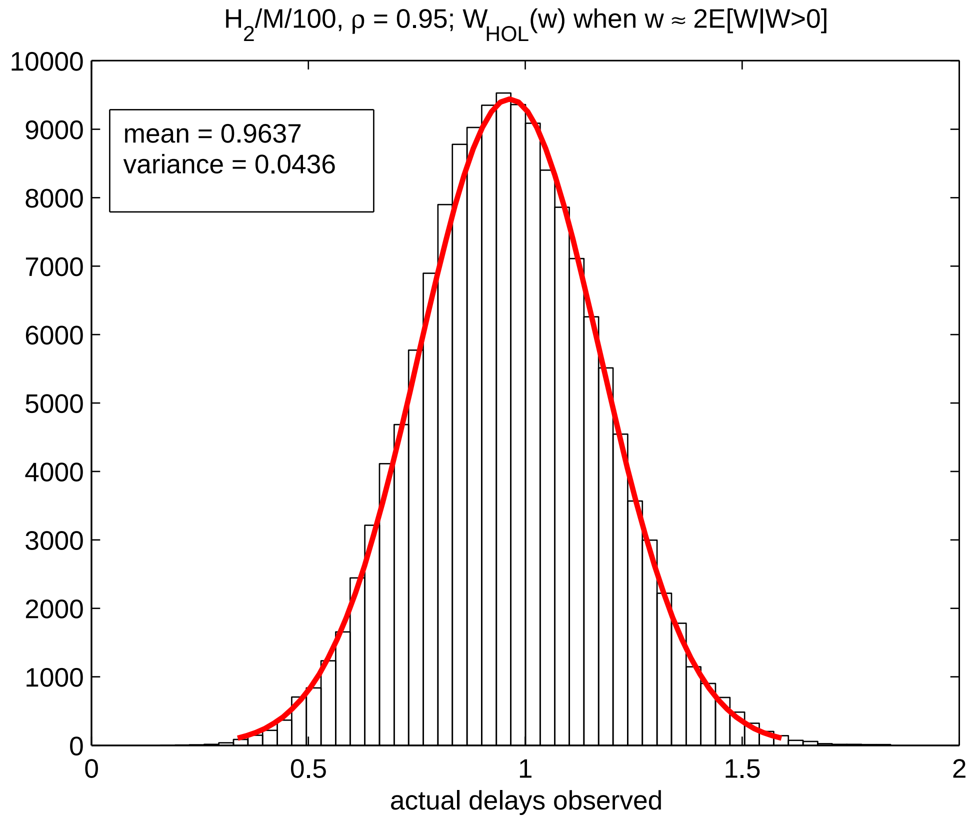

Figure 72: Distribution of $W_{HOL}(w)$ when w $\approx 2E[W|W>0]$ in the $H_2/M/100$ model when $\rho = 0.95$

{HM100_actual

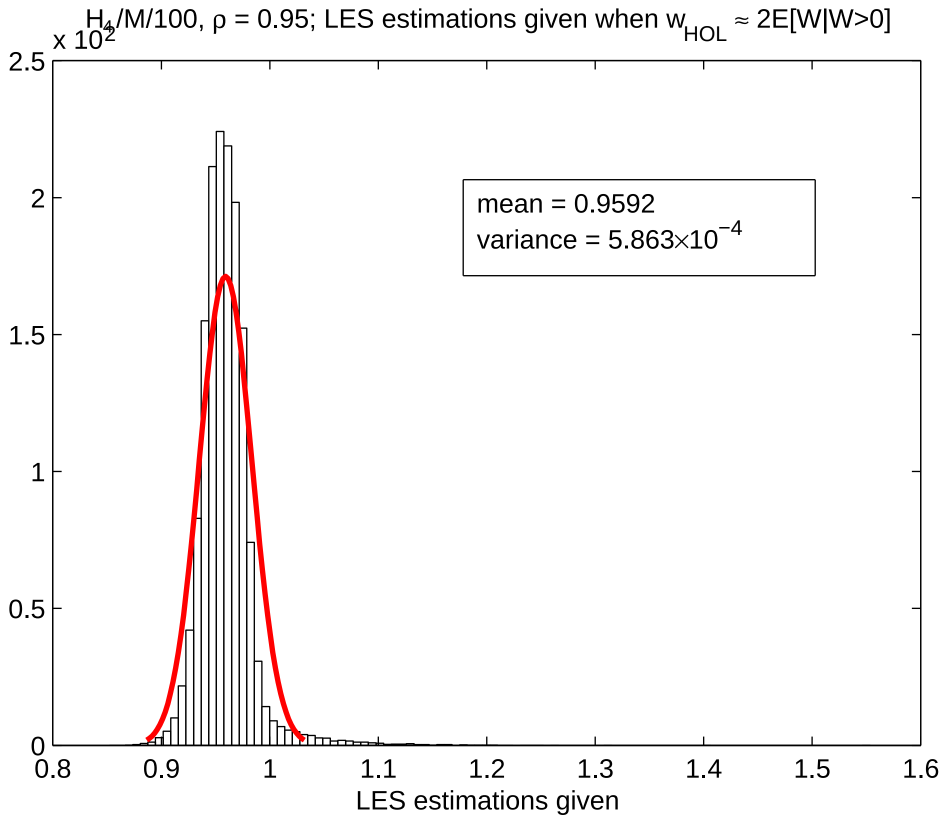

Figure 73: Distribution of the LES delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W>0]$ in the $H_2/M/100$ model when $\rho = 0.95$

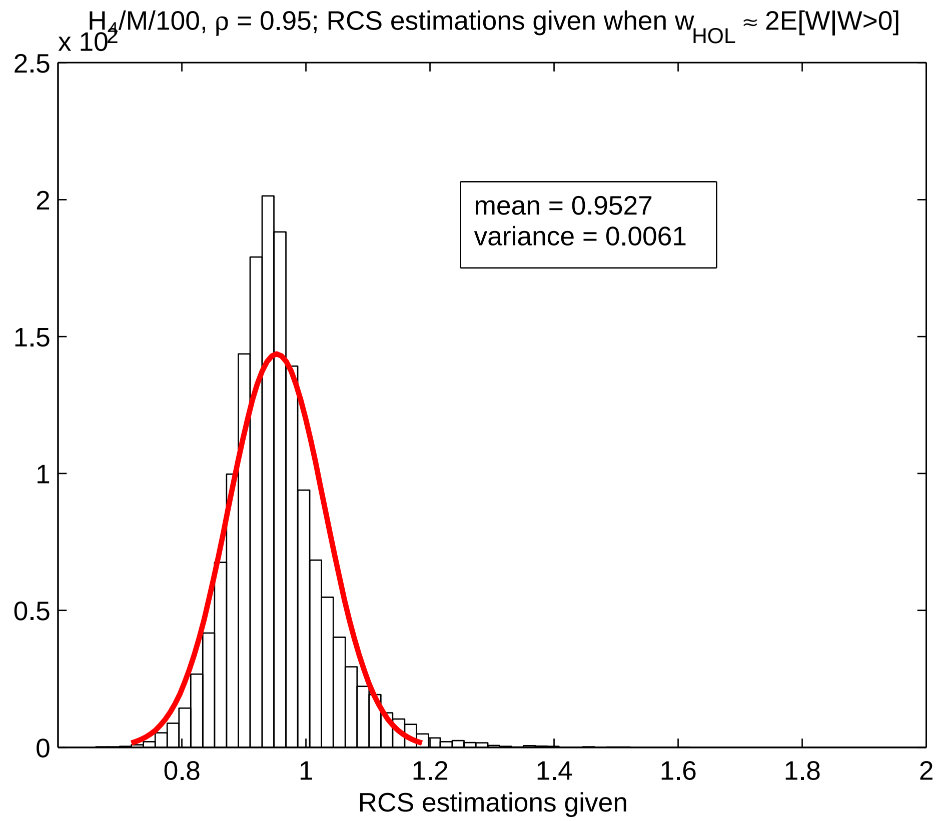

Figure 74: Distribution of the RCS delay estimations given when the HOL estimations, w, are such that w $\approx 2E[W|W > 0]$ in the $H_2/M/100$ model when $\rho = 0.95$

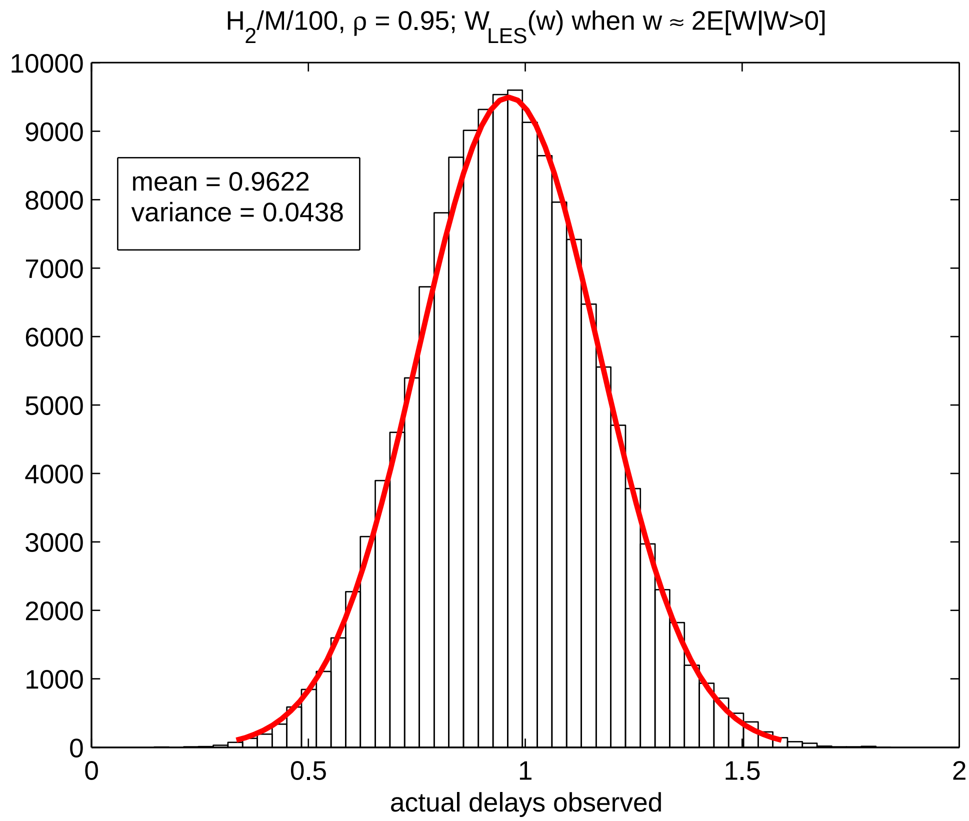

Figure 75: Distribution of the actual delays observed when the LES delay estimation, w, is such that: w $\approx 2E[W|W > 0]$ in the $H_2/M/100$ model when $\rho = 0.95$

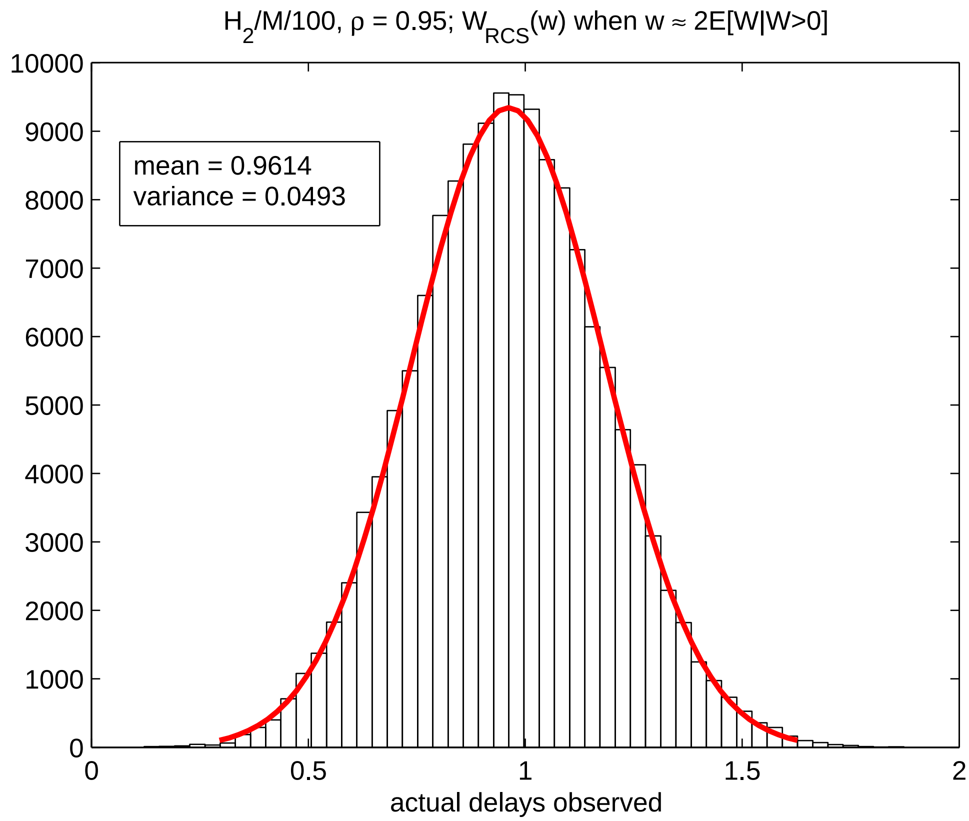

Figure 76: Distribution of the actual delays observed when the RCS delay estimation, w, is such that: w $\approx 2E[W|W > 0]$ in the $H_2/M/100$ model when $\rho = 0.95$

References

Abate, J. and W. Whitt. 1994. A heavy-traffic expansion for asymptotic decay rates of tail probabilities in multichannel queues. *Operations Res. Letters* 15, 223–230.

Abramowitz, M. and I. A. Stegun. 1972. *Handbook of Mathematical Functions*, National Bureau of Standards, U. S. Dept. of Commerce, Washington, D.C.

Armony, M., N. Shimkin and W. Whitt. 2006. The impact of delay announcements in many-server queues with abandonments. Submitted for publication.
Available at http://columbia.edu/∼ww2040.

Asmussen, S. 2003. *Applied Probability and Queues*, second edition, Springer, New York.

Avramidis, A. N., A. Deslauriers and P. L'Ecuyer. 2004. Modeling daily arrivals to a telephone call center. *Management Sci.* 50, 896–908.

Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn and L. Zhao. 2005. Statistical analysis of a telephone call center: a queueing-science perspective. *J. Amer. Statist. Assoc.* 100, 36–50.

Choudhury, G. L. and W. Whitt. 1994. Heavy-traffic asymptotic expansions for the asymptotic decay rates in the BMAP/G/1 Queue. *Stochastic Models* 10, 453–498.

Coates, M., A. O. Hero, III, R. Nowak and B. Yu. 2002. Internet tomography. *IEEE Signal Processing Magazine* 19, 47–65.

Cooper, R. B. 1981. *Introduction to Queueing Theory*, second edition, North-Holland, New York.

Doytchinov, B., J. Lehoczky and S. Shreve. 2001. Real-time queues in heavy traffic with earliest-deadline-first queue discipline. *Ann. Appl. Probab.* 11, 332-378.

Gans, N., G. Koole and A. Mandelbaum. 2003. Telephone call centers: tutorial, review and research prospects. *Manufacturing and Service Opns. Mgmt.* 5, 79–141.

Glynn, P. W. and W. Whitt. 1989. Indirect estimation via $L = \lambda W$. *Operations Research* 37, 82–103.

Glynn, P. W. and W. Whitt. 1994. Logarithmic Asymptotics for steady-state tail probabilities in a single-server queue. *J. Appl. Prob.* 31A, 131–156 (also called *Studies in Applied Probability, Papers in Honour of Lajos Takcs*, J. Galambos and J. Gani (eds.), Applied Probability Trust, Sheffield, England).

Glynn, P. W. and A. J. Zeevi. 2000. Estimating tail probabilities in queues via extremal statistics. In *Analysis of Communication Networks: Call Centres, Traffic and Performance*, D. R. McDonald and S. R. E. Turner (eds.), Fields Institute Communications 28, American Math. Soc., Providence, RI, 135–158.

Guo, P. and P. Zipkin 2004. Analysis and comparison of queues with different levels of delay information. Duke University.

Halfin, S. and W. Whitt. 1981. Heavy-traffic limits for queues with many exponential servers. *Operations Research* 29, 567–588.

Hassin, R. and M. Haviv. 2003. *To Queue or Not to Queue: Equilibrium Behavior in Queueing Systems*, Kluwer.

Ibrahim, R. E. and W. Whitt. 2006. Real time delay estimation based on delay history in the $GI/M/s$ queue. IEOR Department, Columbia University, New York, NY.

Iglehart, D. L. and W. Whitt. 1970. Multiple channel queues in heavy traffic II: sequences, networks, and batches. *Advances in Applied Probability* 2, 355–369.

Jelenkovic P., A. Mandelbaum A. and P. Momcilovic. 2004. Heavy traffic limits for queues with many deterministic servers. *Queueing Systems* 47, 53–69.

Mandelbaum A., A. Sakov and S. Zeltyn. 2000. Empirical analysis of a call center. Technical Report, Faculty of Industiral Engineering and Management, The Technion, Israel.

Nakibly, E. 2002. *Predicting Waiting Times in Telephone Service Systems*, MS thesis, the Technion, Haifa, Israel.

Neuts, M. F. 1986. The caudal characteristic curve of queues. *Adv. Appl. Probab.* 18, 221–254.

Puhalskii, A. A. and M. I. Reiman. 2000. The multiclass $GI/PH/N$ queue in the Halfin-Whitt regime. *Adv. Appl. Prob.* 32, 564–595.

Reiman, M. I. 1982. The heavy traffic diffusion approximation for sojourn times in Jackson networks. In *Applied Probability – Computer Science, the Interface, II*, R. L. Disney and T. J. Ott (eds.), Birhauser, Boston, 409–422.

Ross, S. M. 1996. *Stochastic Processes*, second edition, Wiley, New York.

Shimkin, N. and A. Mandelbaum. 2004. Rational abandonment from tele-queues: nonlinear waiting cost with heterogeneous preferences. *Queueing Systems* 47, 117–146.

Smith, W. L. 1953. On the distribution of queueing times. *Proc. Camb. Phil. Soc.* 49, 449–461.

Ward, A. W. and W. Whitt. 2000. Predicting response times in processor-sharing queues. In *Analysis of Communication Networks: Call Centres, Traffic and Performance*, D. R. McDonald and S. R. E. Turner (eds.), Fields Institute Communications 28, American Math. Society, Providence, RI, 1-29.

Whitt, W. 1982. Approximating a point process by a renewal process: two basic methods. *Operations Research* 30, 125–147.

Whitt, W. 1984. On approximations for queues, I: extremal distributions. *AT&T Bell Lab. Tech. J.* 63, 115–138.

Whitt, W. 1989. Planning queueing simulations. *Management Sc.* 35, 1341–1366.

Whitt, W. 1999. Predicting queueing delays. *Management Sci.* 45, 870–888.

Whitt, W. 2002. *Stochastic-Process Limits*, Springer, New York.

Whitt, W. 2004. A diffusion approximation for the G/GI/n/m queue. *Operations Research* 52, 922–941.

Whitt, W. 2005. Heavy-traffic limits for the G/H2*/n/m queue. *Math. Oper. Res.* 30, 1–27.