
Qian-Yu Tang, Han-Fu Chen, and Zeng-Jin Han

Abstract— In this paper the Perturbation-Analysis-Robbins–Monro-Single-Run algorithm is applied to estimating the optimal parameter of a performance measure for the GI/G/1 queueing systems, where the algorithm is updated after every fixed-length observation period. Our aim is to analyze the limiting behavior of the algorithm. The almost sure convergence rate of the algorithm is proved in [5], [6], and [13], and [20]. The proofs of weak convergence and convergence convergence as reported in the empirical studies [17]–[19]. However, noise typically constitutes a martingale difference sequence (m.d.s.), be performed after every fixed-length observation period. In the first after the observation of one or two regenerative cycles and may also take advantage of the regenerative structure of the system, but the implementation of the algorithm does not depend on the regenerative structure. Thus, the PARMSR algorithm with fixed-length observation period may be applicable to much more general SDES’s.

The rest of the paper is organized as follows. For simplicity of exposition, we give the PARMSR algorithm, updated every customer, in Section II. The proofs of the main results are presented in Section III. In Section IV the obtained results are extended to the case where parameter updates are performed after every fixed number of customers per period. Finally, a concluding remark is given in Section V.

I. INTRODUCTION

Perturbation analysis (PA), since introduced by Ho et al. [11], has been widely studied in the literature on stochastic discrete-event systems (SDES’s); see, for example, Ho and Cao [10], Glasserman [8], and the references therein. Roughly speaking, PA is a method for estimating derivatives of performance measures with respect to system parameters from a single sample path of an SDES, where analytic formulas of the performance measures are only available for a limited class of SDES’s. Combining the PA technique with stochastic approximation algorithms leads to the so-called “single-run optimization” algorithm. When the Robbins–Monro (RM) algorithm is applied, the resulting algorithm is called the “Perturbation-Analysis-Robbins–Monro-Single-Run” (PARMSR) algorithm in Suri and Leung [18] and Suri [17]. The PARMSR algorithm is used for seeking the optimal parameter of a performance measure based on a single sample path of the system.

For PARMSR algorithms, the parameter updates may be performed after the observation of one or two regenerative cycles and may also be performed after every fixed-length observation period. In the first case, convergence analysis is relatively simple, since the observation noise typically constitutes a martingale difference sequence (m.d.s.), and the standard stochastic approximation results are applicable; see, e.g., [4] and [7]. In this case, the parameter updates may be infrequent, since the regenerative cycles may be long in a high-load system as well as in a queueing network with many nodes. In the second case, the PARMSR algorithm has a relatively fast rate of convergence as reported in the empirical studies [17]–[19]. However, its convergence was not proved until recently; see, [5], [6], [12], [13], and [20]. The proofs of weak convergence and convergence in probability of the algorithm are provided in [12] and [13] (with numerical experiments in a companion paper [14]), respectively. The almost sure convergence of the algorithm is proved in [5], [6], and [20].

As pointed out in [4], [5], and [13], the analysis of the convergence rate of the PARMSR algorithm with fixed-length observation period is an interesting and difficult problem and has been lacking. The difficulties lie in the fact that the standard conditions for the convergence rate established in the literature on stochastic approximation are not quite verifiable in the special context of SDES’s. In this paper, we establish the convergence rates of the PARMSR algorithms with a fixed-length observation period for the GI/G/1 queueing systems. It is shown that the convergence rates of the PARMSR algorithms depend on the second derivative of the performance measure at the optimal point. It is worth noticing that our analysis of the convergence rates takes advantage of the regenerative structure of the system, but the implementation of the algorithm does not depend on the regenerative structure. Thus, the PARMSR algorithm with fixed-length observation period may be applicable to much more general SDES’s.

The rest of the paper is organized as follows. For simplicity of exposition, we give the PARMSR algorithm, updated every customer, in Section II. The proofs of the main results are presented in Section III. In Section IV the obtained results are extended to the case where parameter updates are performed after every fixed number of customers per period. Finally, a concluding remark is given in Section V.

II. THE PARMSR ALGORITHM UPDATED EVERY CUSTOMER FOR THE GI/G/1 QUEUE

Let us consider a special regenerative system, the GI/G/1 queueing system, with service in order of arrival, where the ith customer that enters the system is denoted by \( C_i \), \( i \geq 1 \). The interarrival times \( \{ A_n, n \geq 1 \} \) and the service times \( \{ X_n(\theta), n \geq 1 \} \) are i.i.d. sequences and are mutually independent with the first moments \( E A_1 = \frac{1}{\lambda} \) and \( E X_n(\theta) = \pi(\theta) \), respectively, where \( \theta \) is a decision parameter which can be adjusted. The traffic intensity is denoted by \( \rho(\theta) = \frac{\lambda}{\mu} \pi(\theta) \). Throughout the paper, we assume \( \rho(\theta) < 1, \forall \theta \in D \), where \( D \) is a compact set.

Let \( T_n(\theta) \) be the system time of the customer \( C_n, \forall n \geq 1 \). We discuss the performance measures of the type

\[
J(\theta) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} J(T_i(\theta), \theta)
\]

where \( J(t, \theta) \) is a differentiable function with respect to \( (t, \theta), \forall t \geq 0, \theta \in D \). For example, we can choose \( J(t, \theta) = t + C(\theta) \), where \( C(\theta) \) is a known function; see, [4], [5], [7], [13], [18]–[20], etc.

Formally, our problem under consideration is to search \( \theta_0^* \) such that \( J(\theta_0^*) = \max_{\theta \in D} J(\theta) \).

We now define our recursive procedure. Set \( f(\theta) = dJ(\theta)/d\theta \). We use the following projected RM algorithm to update the parameter estimate \( \theta_{n+1} \):

\[
\hat{\theta}_{n+1} = \hat{\theta}_n - a_n f_{n+1}
\]

\[
\theta_{n+1} = \hat{\theta}_{n+1} + a_{n+1} I_{[\theta_{n+1} < \hat{\theta}_{n+1}]} + a_{n+1} I_{[\theta_{n+1} \geq \hat{\theta}_{n+1}]}
\]

(1)

where \( f_{n+1} \) is the \( (n+1) \)th step derivative estimate by infinitesimal PA. Let \( Q_n \) denote the queue length at the time instant when the customer \( C_n \) leaves the server. By the perturbation propagation rule, the \( (n+1) \)th step estimate for \( dT_{n+1} \) at \( \theta \) is given by

\[
\alpha_{n+1} = a_n I_{[Q_n \geq 1]} + \frac{dx_{n+1}(\theta_n)}{d\theta}
\]

(2)

where \( dx_{n}(\theta)/d\theta, \forall n \geq 1 \) can be computed by the “inversion” method. Let \( F(\theta, x) \) be the distribution function of \( x_{n}(\theta) \) and let
\{u_{n}, i \geq 1\} be an i.i.d. sequence with a uniform distribution on (0, 1]. Define \(\xi_s \sim \P(\theta, u_s) = \inf \{x : F(\theta, x) > u_s\}, \forall i \geq 1\), from which the derivatives \(dx_s(\theta)/d\theta, \forall i \geq 1\) can be obtained; see, e.g., [8], [10], and [19]. Thus the \((n+1)\)th step estimate for \(\hat{\theta}(\cdot)\) is given by

\[
 f_{n+1} = J_1(T_{n+1}, \theta_n) + J_0(T_{n+1}, \theta_n)
\]

(3)

where \(T_{n+1}\) is the system time of the customer \(\gamma_{n+1}\), and \(J_1(\cdot, \cdot)\) and \(J_0(\cdot, \cdot)\) denote the partial derivatives of \(J(\cdot, \cdot)\) with respect to its first and second component, respectively. Then we obtain the PARMSR algorithm updated every customer by combining (1) with (2) and (3). The observation noise is expressed as

\[
 \varepsilon_{n+1} = f_{n+1} - f(\theta_n).
\]

(4)

For our results, let us introduce the following conditions. Notice that A3) and A4) are exclusive.

A1) \(0 < \alpha_n \leq \tilde{\alpha}_n \rightarrow^\alpha \nu, \forall \nu \in \left(\frac{1}{2}, 1\right], \forall \alpha_n \geq 1; \Sigma_{n=1}^\alpha = \infty \leq \alpha_n^{-1} \rightarrow^\alpha n^{-\alpha}, n \rightarrow \infty, \alpha > 0\).
A2) \(f(\theta)\) has a unique root \(\theta^0 \in (a, b) \subset \mathbb{D}, J(\theta^0) = 0\).
A3) As \(\theta \rightarrow \theta^0, f(\theta)\) can be expressed as \(f(\theta) = M_1(\theta - \theta^0) + \Delta(\theta)\), where \(\Delta(\theta) = O(\|\theta - \theta^0\|^p)\) as \(\theta \rightarrow \theta^0, M_1 \neq 0\), and \(\theta \in [0, 1/2]\) is a constant.
A4) \(f(\theta) = M_2(\theta - \theta^0)^{\alpha} \rightarrow \theta^0 \Rightarrow r(\theta), r(\theta) = o(\|\theta - \theta^0\|^{1+\gamma})\) as \(\theta \rightarrow \theta^0, M_2 > 0, \gamma > 0\).
A5) \(\Sigma_{n=1}^{\alpha_n}(a_n \rightarrow \alpha_n^{-1}) = \infty\).
A6) There are constants \(B_1, B_3, \mu_0, \mu_2\) such that

\[
 \max \left\{||J_0(\theta, \theta)\|, |J_1(\theta, \theta)|\right\} \leq B_1 + B_2 t^{\alpha_n}.
\]

(5)

\[
 \frac{dx_2(\theta)}{d\theta} = \Phi(\theta, x_2(\theta))
\]

\[
 |\Phi(\theta, x)| \leq \sum_{j=0}^{\nu} b_{j} x_{\nu}, \quad \forall \theta \in [a, b], x \geq 0
\]

where \(E x_{\nu}^{\alpha_n} < \infty\).

A8) We have the following.

a) There are two positive constants \(\beta_0\) and \(\mu\) such that \(P(t_1 \leq A_1 \leq t_1 + t_2) \leq \beta_0 t^{2\alpha}, \forall i, t_1 \geq 0\).

b) \(E A_1^{\alpha_n} < \infty\).

c) \(\sup_{\theta \in [a, b]} E x_2(\theta) < \infty\), where \(\xi = \alpha_0 \max \{2(3\nu + 4p_0 + \nu_1 + \mu_2), 2p_2(p + p_0), 2g/(p_1 - 1)(1+p_0)\}, p_1 > 1, p_2 > 1, g \in [0, \mu], (iv) \delta \in [0, 1 - (1/2\nu), \nu/(1 - \delta) + (1/p_1)(1 - (1/p_2)\xi)] > 1\).

Conditions A1) and A5) on the step sizes are standard; for example, we can choose \(a_n \sim \tilde{\alpha}_n^{-\alpha}, \forall \nu \in (1/2, 1]\). Since our main concern is with the convergence rate of the algorithm, Condition A2) is reasonable, i.e., \(J(\theta)\) has a unique minima \(\theta^0\) in \((a, b)\). Condition A3) requires that \(J(\theta)\) has positive second derivative at \(\theta^0\), while Condition A4) says that the second derivative of \(J(\theta)\) at \(\theta^0\) is zero. It will be shown that, roughly speaking, the convergence rate of \(\|\theta_n - \theta^0\| = o(a_n^{-\alpha})\) and \(O((\log a_n^{-\alpha})^{1/\gamma})\), respectively, under Conditions A3) and A4).

The bounds in Condition A6) are not essential for the convergence analysis, since \(\mu_0, \mu_1, \mu_2\) are arbitrary. If \(J(t, \theta) = t + C(\theta)\), then \(\mu_0 = \mu_1 = \mu_2 = 0\) and \(B_2 = B_3 = 0\). This performance function has been widely discussed; see, e.g., [4]–[5], [7], and [18]–[20]. Additionally, we need the Lipschitz condition on \(f(\theta)\) in A6). Assumption A7) holds if, for \(F(\theta, x)\), either \(\theta\) is a location parameter or \(\theta\) is a scale parameter. In this case, we can set \(p = 1, W_0 = 0\); see [19]. If \(A_1\) has a bounded probability density function, then we can choose \(\mu = 1\) in (A8)-a). Comparing it with that used in [5] and [6], where the distribution of \(A_1\) is assumed to have a bounded hazard rate, our condition is rather weak. Since the distribution of \(A_1\) is independent of \(\theta\), the convergence of the PARMSR algorithm should not depend on the distribution of \(A_1\). This is proved in [20]. Some moment conditions on the service times and the interarrival times are required in Condition A8)-b) and c).

The main results of this paper are as follows.

Theorem 2.1: Assume that A1), A2), and A6)–A8) hold, then \(\theta_n \rightarrow \theta^0, \forall \nu \in (1/2, 1]\).

Theorem 2.2: Let A1)–A3) and A6)–A8) hold. Then \(\|\theta_n - \theta^0\| = o(a_n^{-\alpha})\), for some \(\delta \in [0, 1 - (1/2\nu)]\) such that Conditions A3) and A8)-c) are satisfied.

Theorem 2.3: If A1), A2), and A4)–A8) hold, then

\[
 (\log a_n^{-\alpha})^{1/\gamma} \|\theta_n - \theta^0\| \rightarrow^\alpha \left(\frac{\alpha}{M_2^2}\right)^{1/\gamma}, \quad \text{a.s.}
\]

III. THE PROOFS OF THE MAIN RESULTS

With the PARMSR algorithm applied, we denote the number of customers served in the \((m + 1)\)th busy period (BP) by \(N_{m+1}\). Then we have

\[
 N_{m+1} = \inf \left\{t : \sum_{j=0}^{\nu}(A_k(x_{k+j} - x_{k+j-1}(\theta_{m+j-1})) > 0\right\}
\]

(7)

where \(k_m \triangleq \sum_{i=1}^{m} \eta_i, \forall m \geq 1, k_0 = 0\). By A7), \(dx_j(\theta)/d\theta\) does not change sign, \(\forall i \geq 1, \theta \in [a, b]\). Without loss of generality, we assume that \(dx_j(\theta)/d\theta > 0, \forall i \geq 1, \theta \in [a, b]\) henceforth. If \(\theta_{m+k+1}, \theta_{m+k+2}, \ldots\) are replaced by \(\theta\), we denote the number of customers in the \((m + 1)\)th BP by \(N_{m+1}(\theta)\). Denote the length of the \((m + 1)\)th BP and the \((m + 1)\)th idle period by \(I_{m+1}\) and \(I_{m+1}\), respectively, and by \(L_{m+1}(\theta)\) and \(L_{m+1}(\theta)\) if the service parameters are fixed at \(\theta\) throughout the BP. More precisely, we have

\[
 L_{m+1}(\theta) = \sum_{i=0}^{\nu}(x_{k_{m+i}}(\theta_{m+i} - \theta_{m+i} - 1))
\]

\[
 L_{m+1} = \sum_{i=0}^{\nu}(x_{k_{m+i}}(\theta_{m+i} - \theta_{m+i} - 1))
\]

\[
 I_{m+1}(\theta) = \sum_{i=0}^{\nu}(x_{k_{m+i}} - x_{k_{m+i}}(\theta_{m+i} + 1))
\]

Before proving our main results, we need several lemmas.
Lemma 3.1: Let A1) and A6)–A8) hold. Then for any fixed \( m', \ell' \), and all \( \delta \in [0, 1 - (1/2\nu)] \), we have

\[
\sum_{m=-\infty}^{k_{m+\ell'}+\ell'} a^1_{m+i} \frac{\delta}{2} \rightarrow 0, \text{ a.s.}
\]

Proof: For all \( i \geq 0 \), \( 1 \leq j \leq \eta_{m+i+1} \), by (5) and (7) it follows that

\[
T_{k_{m+i}+j} = \sum_{s=1}^{j} x_{m+i+j} (t_{k_{m+i}+j} - 1) = \sum_{s=1}^{j-1} A_{m+i+j+1} \leq L_{m+i+1}(b) 
\]

(8)

\[
|\alpha_{m+i+j}| \leq \frac{1}{\max_{\theta \in [a, b]} |f(\theta)|} \int_{r_{m+i}+1}^{r_{m+i}+\ell'} |f(\theta)| \sum_{m=0}^{p} b_s(L_{m+i+1}(b))^r 
\]

(9)

Noticing that \( \alpha_{m+i} \leq \eta m^{-\nu}, \forall m \geq 1 \), by (2), (3), (8), (9), and Condition A6) we get

\[
\sum_{m=-\infty}^{k_{m+\ell'}+\ell'} a^1_{m+i} \frac{\delta}{2} \rightarrow 0, \text{ a.s.}
\]

(10)

By A8) and c) it follows that (see, e.g., [9] and [20])

\[
E \int_{m}^{\infty} f(\theta) \sum_{m=0}^{p} b_s(L_{m+i+1}(b))^r 
\]

(11)

By (11) and [20, Lemma 2], from (10) the assertion of Lemma 3.1 follows immediately.

Lemma 3.2: Suppose that one of the following conditions is satisfied.

1) \( \forall t \geq 0, t_i \geq 0, P\{t_i \leq A_1 \leq t_i + t\} \leq \beta_0 t^r \); 2) \( \forall t \geq 0, t_i \geq 0, P\{t_i \leq x_1(\theta) \leq t_i + t\} \leq \beta_0 t^r \). Then

\[
E \frac{1}{\max_{\theta \in [a, b]} |f(\theta)|} \sum_{m=0}^{p} b_s(L_{m+i+1}(b))^r 
\]

(12)

Since \( \{Y_i, i \geq 1\} \) is an i.i.d. sequence, by (12) we get

\[
P\{I_1(\theta) \leq t\} = P\left\{ \sum_{i=1}^{\infty} Y_i \leq t \right\} = \sum_{k=1}^{\infty} P\left\{ \sum_{i=1}^{k} Y_i \leq t, \eta_1(\theta) = k \right\} 
\]

(13)

By (13) we then derive

\[
EI_{1}^{-\nu}(\theta) = \int_{0}^{\infty} t^{-\nu-1} P\{I_1(\theta) < t\} dt 
\]

\[
= \int_{0}^{\infty} t^{-\nu-1} P\{I_1(\theta) < t\} dt + \int_{t}^{\infty} t^{-\nu-1} P\{I_1(\theta) < t\} dt 
\]

(14)

2): The proof is similar to that of 1).

The proof of Lemma 3.2 goes through if, instead, we use the conditional probability version.

Corollary 3.1: If Condition A8)-a) holds, then \( E(I_{m+1} F_{k_{m}}) \leq \beta_1, \forall m \geq 1 \) \( \in (0, \mu) \) where the filtration \( F_{k_{m}} \) is defined by

\[
F_{k_{m}} = \sigma \{B: \sigma \in \{A_j, u_j, \forall j \geq 1\} \}
\]

and

\[
B[k_{m} = i] \in \sigma \{A_j, u_j, \forall j \geq 1\} \forall i \geq 1 \}
\]

Lemma 3.3: Let A6)–A8) hold. Then there exist constants \( \beta_2 > 0 \) and \( \gamma \in (0, q) \) such that

\[
P\{\eta_{m+1} \neq \eta_{m+1}(\theta, k_{m}) | F_{k_{m}} \} \leq \beta_2 a^{\gamma}_{m+1}, \forall m \geq 1 \}
\]

Proof: By Condition A6), (8), and (9), from (1)–(3) it follows that

\[
|\theta_{k_{m+i}} - \theta_{k_{m}}| \leq \sum_{j=0}^{i} a_{k_{m+i}-1} |I_{k_{m+i}}| 
\]

(15)

where

\[
W_{m+1} \triangleq \eta_{m+1}(b)(B_{1} + B_{2} L_{m+1}(b)) 
\]

(16)
By the Hölder inequality and Lemma 3.2, it follows from (7) that
\[
\begin{align*}
\text{P}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|F_{k_m}\} & \leq P_{\theta_{k_m}}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})\} \\
& \leq P_{\theta_{k_m}}\left\{ I_{m+1}(\theta_{k_m}) \leq \sum_{i=1}^{m+1} (x_{k_m+i-1} - x_{k_m+i})(\theta_{k_m}) \right\} \\
& \leq a_{k_m} W^{(1)}_{m+1} \tag{17}
\end{align*}
\]
where
\[
W^{(1)}_{m+1} = W_{m+1} \sum_{i=0}^{p} b_i E^{(2)}_{m+1}(b).
\tag{18}
\]

By the Hőlder inequality and Lemma 3.2, it follows from (7) that
\[
\text{P}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|F_{k_m}\} \leq P_{\theta_{k_m}}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})\} \\
\leq P_{\theta_{k_m}}\left\{ I_{m+1}(\theta_{k_m}) \leq \sum_{i=1}^{m+1} (x_{k_m+i-1} - x_{k_m+i})(\theta_{k_m}) \right\} \\
\leq a_{k_m} W_{m+1} \tag{17}
\]
where
\[
I_{m+1}(\theta_{k_m}) \leq \sum_{i=1}^{m+1} (x_{k_m+i-1} - x_{k_m+i})(\theta_{k_m})
\]

which yields
\[
\beta_2 < \infty.
\]
Similarly, we have
\[
\text{P}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|F_{k_m}\} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

By (8), (9), and Condition A6) it follows from (21) that
\[
\text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq \text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

which yields
\[
\beta_2 < \infty.
\]

Similarly, we have
\[
\text{P}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|F_{k_m}\} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

By (8), (9), and Condition A6) it follows from (21) that
\[
\text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq \text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

which yields
\[
\beta_2 < \infty.
\]

Similarly, we have
\[
\text{P}\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|F_{k_m}\} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

By (8), (9), and Condition A6) it follows from (21) that
\[
\text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq \text{E}\{ I_{m+1}(\theta_{k_m}) \} \leq a_{k_m}^{\beta_2} W^{(2)}_{m+1} \tag{21}
\]

which yields
\[
\beta_2 < \infty.
\]
from which we get
\[
\sum_{m=1}^{\infty} a_{k_m}^2(1-\delta) \leq \sum_{m=1}^{\infty} \left( \sum_{i=1}^{\epsilon_{m+1}(B_{k_m})} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) \right)^2 \mathcal{F}_{k_m} < \infty.
\]

Then by the local convergence theorem of martingales (see, e.g., [16]), the second term on the right-hand side of the equality in (19) converges a.s.

3): By A6) and (15), we have
\[
\left| \alpha_{k_{m+1}} - \alpha_{k_{m+1}}(\theta_{k_{m+1}}) \right| \leq B_5 \alpha_{k_m} \eta_{m+1}(b) W_{m+1}.
\]

Then the third term on the right-hand side of the equality in (19) converges a.s., by the local convergence theorem of martingales (see, e.g., [16]).

4): Using A6), A7), and (15), it follows from (20) that
\[
\left| \alpha_{k_{m+1}} - \alpha_{k_{m+1}}(\theta_{k_{m+1}}) \right| \leq W_0 \alpha_{m+1}(b) a_{k_m} W_{m+1}
\]

and similar to (17) it is seen that
\[
\left| \frac{1}{m} \sum_{j=1}^{m} x_{k_{m+1}}(\theta_{k_{m+1}}) - x_{k_{m+1}}(\theta_{k_{m+1}}) \right| \leq a_{k_m} W_{m+1}^{(1)}
\]

where \( W_{m+1}^{(1)} \) is defined by (18). By (3) and (20) we have
\[
\left| f_{k_{m+1}} - f_{k_{m+1}}(\theta_{k_{m+1}}) \right| \leq \left| J_1(T_{k_{m+1}}, \theta_{k_{m+1}}) - J_1(T_{k_{m+1}}, \theta_{k_{m+1}}) \right|
\]

where
\[
\lim_{m \to \infty} a_{k_m} W_{m+1}^{(1)} = 0.
\]

Using (11) and the Schwarz inequality, we see that
\[
\sup_{m} E \left\{ \eta_{m+1}(b) W_{m+1} | \mathcal{F}_{k_m} \right\} < \infty
\]

which in conjunction with the local convergence theorem of martingales yields that the fourth term on the right-hand side of the equality in (19) converges a.s.

5): Similar to (22) we can prove
\[
\sum_{m=1}^{\infty} \sum_{i=1}^{\epsilon_{m+1}(B_{k_m})} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) I_{\{\eta_{m+1}(\theta_{k_m}) > \eta_{m+1}(\theta_{k_m})\}} \leq W_{m+1}^{(3)}
\]

where
\[
W_{m+1}^{(3)} = \eta_{m+1}(b) W_{m+1}^{(3)}.
\]

By Lemma 3.3 and the Hölder inequality, we derive
\[
E \left\{ a_{k_m}^{1-\delta} \sum_{i=1}^{\epsilon_{m+1}(B_{k_m})} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) \right| \mathcal{F}_{k_m} \}
\]

\[
\leq a_{k_m}^{1-\delta} \left\{ E W_{m+1}^{(3)} \right\}^{1/2} P_{\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m}) \}} \mathcal{F}_{k_m}^{1/2}
\]

\[
\leq a_{k_m}^{1-\delta} \left\{ E W_{m+1}^{(3)} \right\}^{1/2} \left\{ E W_{m+1}^{(3)} \right\}^{1/2}
\]

where \( (1/p_2) + (1/q_2) = 1 \). Choosing the appropriate \( p_1, q_2 \) such that \( p_1(1-\delta) + (1/p_1 q_2) > 1 \), then by the local convergence theorem of martingales we obtain that the fifth term on the right-hand side of the equality in (19) converges a.s.

Similarly, we can prove that the last term on the right-hand side of the equality in (19) converges a.s. Thus we conclude the proof of the lemma.

Lemma 3.5: If A1) and A6)–A8) hold, then \( \sum_{m=1}^{\infty} a_{m}^{1-\delta} \epsilon_{m+1} \) converges a.s., for all \( \delta \in [0, 1) \).

Proof: Let the sample path be fixed. Define \( M(i) = \sup \{ j: \epsilon_j \leq i \} \). \( \forall i \geq 1 \), \( \forall i > 0 \), by Lemmas 3.1 and 3.4 there exists \( n_0 \) such that for all \( n_2 > n_1 \geq n_0 \) we have
\[
\left\{ \sum_{i=1}^{n_2} a_i^{1-\delta} \epsilon_{i+1} \right\}
\]

\[
= \left\{ \frac{1}{m(n_2)} \sum_{m=1}^{m(n_2)-1} \sum_{i=1}^{n_2} a_i^{1-\delta} \epsilon_{i+1} \right\}
\]

\[
\leq \epsilon.
\]

This completes the proof of the lemma.

Proof of Theorems 2.1–2.3: Theorem 2.1 follows easily from Lemma 3.5 and [2, Th. 3.1]. By Lemma 3.5 and [3, Th. 3.2.1] we obtain Theorem 2.2. Finally, by Lemma 3.5 and [1, Th. 3] we have Theorem 2.3.

IV. THE PARMER ALGORITHM UPDATED EVERY \( L \)-CUSTOMERS PERIOD FOR THE GI/G/1 QUEUE

We continue considering the GI/G/1 queueing system. The obtained results can easily be extended to the case where parameter updates are performed after \( L \) customers depart from the server. Let \( f_{n+1} \) be the \( (n+1) \)th step estimate for \( f(\theta) \). Instead of (2) and (3), we now have
\[
f_{n+1} = \frac{1}{L} \sum_{i=1}^{L} J_i(T_{n+i}, \theta_n) \beta_{n+1,i} + \tilde{J}_i(T_{n+i}, \theta_n)
\]

where \( \beta_{n+1,i} \), \( 1 \leq i \leq L \) are defined by
\[
\beta_{n+1,i} = \beta_{n+i,1} \beta_{n+i+1,1} + \frac{d \tau_{n+i+1}(\theta_n)}{d \theta}
\]

\[
0 \leq i \leq L - 1.
\]
The PARMSR algorithm updated every $L$-customers period is composed of (1), (28), and (29); see, e.g., [18]–[20].

**Theorem 4.1**: The assertions of Theorem 2.1–2.3 hold in the present setting, if the conditions of the theorems are satisfied, respectively.

**Proof**: The key step is to verify that $\sum_{\infty} a_n^{-1} \frac{d}{d\theta} \theta_n$, converges, a.s. The observation noise of the algorithm is (4), where $f_{n+1}$ is defined by (28). Along the same lines as in [20], we first introduce the following notations:

$$
\mu(n) = \frac{1}{L} \
\bar{\theta}_n = \theta_{\mu(n)} \
\tilde{a}_n = a_{\mu(n)}
$$

$\beta(n+1, j) = \beta(n, j)$

$$
\tau_n = J_{n}(T_n, \bar{\theta}_{n-1}) \beta_{n, j} + J_{n}(T_n, \bar{\theta}_{n-1}) - f(\bar{\theta}_{n-1})
$$

$\forall n \geq 1$.

By (29) and (31) we have

$$
\beta_{n+1} = \beta_n I_{[\log_2 n + 1]} + \frac{d}{d\theta} \bar{\theta}_n, \quad \forall n \geq 0.
$$

From (28) and (29) it is derived that

$$
\sum_{n=1}^{\infty} a_n^{-1} \frac{d}{d\theta} \tau_n = \sum_{n=1}^{\infty} a_n^{-1} (f_{n+1} - f(\theta_n))
$$

$$
= \frac{1}{L} \sum_{n=0}^{\infty} a_n^{-1} \tau_n
$$

which implies that the almost sure convergence of $\sum_{n=1}^{\infty} a_n^{-1} \frac{d}{d\theta} \tau_n$ is equivalent to the almost sure convergence of $\sum_{n=1}^{\infty} a_n^{-1} \tau_n$.

The proof of the convergence of $\sum_{n=1}^{\infty} a_n^{-1} \tau_n$ works the same way as in Lemma 3.1 and Lemmas 3.3–3.5 if, instead, we replace $\alpha_n, \bar{\theta}_n, \tau_n, \tilde{a}_n$, respectively. Details are omitted, for the brevity of the paper.

**V. CONCLUDING REMARK**

We have established the convergence rates of the PARMSR algorithm with fixed-length observation period for the GI/G/1 queueing systems. Along the same lines of the research, more precise convergence results for the PARMSR algorithms, such as a central limit theorem and a law of the iterated logarithm, could be derived.

**REFERENCES**


[12] M. C. Fu and Nikita E. Barabanov

**Improved Upper Bounds for the Mixed Structured Singular Value**

Minyue Fu and Nikita E. Barabanov

**Abstract**—In this paper, we take a new look at the mixed structured singular value problem, a problem of finding important applications in robust stability analysis. Several new upper bounds are proposed using a very simple approach which we call the multiplier approach. These new bounds are convex and computable by using linear matrix inequality (LMI) techniques. We show, most importantly, that these upper bounds are actually lower bounds of a well-known upper bound which involves the so-called D-scaling (for complex perturbations) and G-scaling (for real perturbations).

**Index Terms**—Robust control, robust stability, robustness, structured singular value, uncertain systems.

**I. INTRODUCTION**

This paper addresses the problem of the mixed structured singular value. The notion of structured singular value, or $\mu$ for short, was initially proposed by Doyle [4] for studying the robust stability...