CENTRAL LIMIT THEOREMS FOR STOCHASTIC OPTIMIZATION ALGORITHMS USING INFINITESIMAL PERTURBATION ANALYSIS

Qian-Yu Tang, Pierre L'Ecuyer, and Han-Fu Chen

July 8, 1998

Abstract. Central limit theorems are obtained for the "perturbation analysis Robbins-Monro single run" (PARMSR) optimization algorithm, with updates either after every L customers or after every busy period, in the context of the optimization of a $GI/GI/1$ queue. The PARMSR algorithm is a stochastic approximation (SA) method for the optimization of infinite-horizon models. It is shown that the convergence rate and the asymptotic variance constant of the optimization algorithm, as a function of the total computing budget (i.e., total number of customers), are the same for both updating methods, and independent of L, provided that the step sizes of SA are chosen in the (asymptotically) optimal way for each method.

Résumé. On obtient des théorèmes de la limite centrale pour la solution de l'algorithme "perturbation analysis Robbins-Monro single run" (PARMSR) pour l'optimisation d'une file d'attente $GI/GI/1$, avec mise à jour après chaque L clients ou après chaque cycle regeneratif. L'algorithme PARMSR est une méthode d'approximation stochastique (AS) pour l'optimisation de modèles sur horizon infini. Nos résultats impliquent que le taux de convergence et la constante de variance asymptotique de l'algorithme d'optimisation, en fonction du budget total de calcul (i.e., le nombre total de clients), sont les mêmes pour les deux méthodes de mise à jour, et indépendants de L, pourvu que les longueurs de pas de l'AS soient choisis de manière optimale pour chaque méthode.

Key words. perturbation analysis, stochastic approximation, recursive estimation, queueing theory, central limit theorems

AMS(MOS) Subject classifications. 62L20, 93E12, 93E23, 93E30, 60F05, 60K25

The Work of Q. Y. Tang and P. L'Ecuyer has been supported by NSERC-Canada grants No. OGP0110050 and SMF0169893, and by FCAR-Québec grant No. 93ER1654. The work of H. F. Chen was partially supported by the National Natural Science Foundation of China. We thank the two anonymous referees, whose comments helped improving the presentation of the results in paper.

Qian-Yu Tang and Pierre L'Ecuyer are with Département d'IRO, Université de Montréal, C.P.6128, Succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada

Han-Fu Chen is with the Laboratory of Systems and Control, Institute of Systems Science, Academia Sinica, Beijing 100080, People's Republic of China.

1 Introduction

Since the advent of perturbation analysis (PA), considerable effort has been devoted to studying PA-based stochastic optimization algorithms for discrete event dynamic systems (DEDSs). Ho and Cao [15] proposed an off-line Robbins-Monro (RM)-type stochastic optimization algorithm combined with a PA derivative estimator. On-line or single-run versions were promoted at an early stage in [11, 23, 31, 32, 33], among others. The RM-type stochastic approximation (SA) algorithm combined with PA in a single run has been referred to as the PARMSR (perturbation analysis Robbins-Monro single run) algorithm by Suri [31] and Suri and Leung [32]. The PARMSR algorithm aims at the optimal design of a DEDS based on a single sample path of the system. Convergence of the PARMSR algorithm has been widely studied in the recent years; see, e.g., [5, 6, 7, 10, 18, 20, 22, 34, 35]. These authors have proved weak or almost sure convergence of the algorithm in certain settings and under different sets of conditions. The convergence rate of the PARMSR algorithm has also been studied empirically in [21] and bounds on the convergence rate have been established theoretically in [35], but precise results on the convergence rate and the limiting variance, in the sense of central limit theorems, have still been lacking. The aim of this paper is to provide such central limit theorems.

The PARMSR idea applies when the performance measure of interest is an average over an infinite time horizon, for a DEDS in steady state. Let $\overline{J}(\theta)$ denote that steady-state performance measure, where the control parameter θ can be adjusted on some compact set D. The goal is to find θ^0 such that $\overline{J}(\theta^0)$ is the global minimal value of $\overline{J}(\theta)$ over D. The PARMSR algorithm is an instance of the general SA algorithm and has the form

$$\theta_{n+1} = \Pi_{n+1}(\theta_n - a_n f_{n+1}), \quad \forall\ n \geq 0, \tag{1.1}$$

where θ_n is the parameter value at the nth SA iteration (i.e., the nth estimate for θ^0), θ_0 is the initial value, Π_{n+1} denotes a projection over D, f_{n+1} is a PA estimator of the derivative $d\overline{J}(\theta)/d\theta$ at $\theta = \theta_n$, and $\{a_n,\ n \geq 1\}$ is the sequence of step sizes. In the next section, we make this setup more precise, in the context of a $GI/GI/1$ queueing model.

When implementing the PARMSR algorithm, an important issue is how to choose the observation length between the SA iterations (see, e.g., [11, 16, 21]). One may observe the evolution of the system for a long period of time to compute each f_{n+1}, which means infrequent updates of the control parameter, or observe it for short periods of time, with frequent changes of the control parameter. Although there has been considerable progress in the convergence analysis of the PARMSR algorithm, as cited above, the problem concerning how the limiting behavior of the PARMSR algorithm depends on the value of the observation length between iterations is still open except for the case where the observation length increases unboundedly with the iteration number, which is studied in [22].

In this paper, we prove central limit theorems for the PARMSR algorithms updated either after every L customers or after every busy period (BP), for a $GI/GI/1$ queueing system. The central limit theorems show how the limiting behavior of the PARMSR algorithms depends on the choice of the step-sizes, the value L, and the average number of customers in a BP. Our results indicate that although a larger L yields a smaller limit variance in terms of the number of iterations of the algorithm, the limiting behavior of the PARMSR algorithm is independent of L when the convergence is expressed in terms of the total computing budget (i.e., total number of customers) and the step sizes are chosen optimally for each L. However, we point out that the small-sample (or transient) behavior of the algorithm may be quite different for different values of L, as explained in [31]–[33]. On the other hand, in the numerical experiments of [21], the behavior of the algorithm was approximately the same for a wide range of values L.

At the early stage of the study of the PARMSR method, Fu [10] and Chong and Ramadge [5] proposed PARMSR algorithms updated after every BP or every two BPs. By updating the control parameter at the start of every one or two BPs, they are able to define an unbiased derivative estimator of a rescaled regression function, which is computed at each iteration, so standard SA results can be applied. But a major drawback is that explicit regeneration points must be available and frequent enough (otherwise the control parameter will not be updated often enough). For queueing networks and manufacturing systems, such regeneration points are typically hard to find or very sparse; see, e.g., [1], [25], [28] and [29]. The PARMSR algorithm updated after every fixed-length observation period avoids this problem, so it can be applied to a larger class of DEDSs.

The comparison of the limiting behavior of the PARMSR algorithm updated after every fixed-length observation period and that updated after every BP (or regenerative cycle) is an attractive problem. As pointed out in [6], "Ideally, we would like to derive expressions for the convergence rates of algorithms involving durations between updates, and thereby be able to compare the convergence rate of an algorithm which updates before the service of each customer with that of one which updates before every busy period. However, this seems to be a difficult task." The present paper accomplishes this task for a $GI/GI/1$ queueing model.

Let N_n denote the total number of customers observed (or simulated) for the first n SA iterations, i.e., for the computation of θ_1, θ_2, ..., θ_n. We recall that if there are two positive constants γ and κ^* such that

$$\frac{(N_n)^\gamma(\theta_n - \theta^0)}{\kappa^*} \xrightarrow[n \to \infty]{d} \mathcal{N}(0, 1), \tag{1.2}$$

where $\xrightarrow{d}$ means convergence in distribution and $\mathcal{N}(0,1)$ is a standard normal random variable, then γ and $(\kappa^*)^2$ are referred to as the convergence rate and the limiting variance constant, respectively. When comparing two algorithms, we prefer the one with the largest γ (best convergence rate) and, in case of a tie, the one with the smallest κ^*. This ranking is based on asymptotics, so it reflects what happens for large enough n. The behavior for small n could differ significantly and is much more difficult to analyze.

The rest of the paper is organized as follows. The recursive procedure for the PARMSR algorithm updated after the service of every customer is formulated in Section 2, for the $GI/GI/1$ queueing model, and a central limit theorem is stated for that case. In Section 3, the result is extended to the

PARMSR algorithm updated after every L customers. In Section 4, we compare the limiting behavior of this algorithm with the algorithm that updates after every busy period. Conclusions are given in Section 5. The proofs are collected in the appendix.

2 The PARMSR Algorithm Updated After the Service of Every Customer in a $GI/GI/1$ Queue

Consider a controlled $GI/GI/1$ queueing model with i.i.d. interarrival times $\{v_i,\ i \geq 1\}$ and i.i.d. service times $\{x_i(\theta),\ i \geq 1\}$, where these random variables are mutually independent, and where θ is a decision parameter that can be adjusted. Throughout the paper, $D = [a, b]$ is assumed to be a closed interval on the real line, and the load condition $\rho(\theta) \stackrel{\triangle}{=} \mathrm{E}[x_1(\theta)]/\mathrm{E}[v_1] < 1$ for all $\theta \in D$ is assumed to hold. We could take D as a more general compact subset of the d-dimensional real space and our arguments would still go through, but for simplicity, we stick to the one-dimensional case. Let $T_n(\theta)$ be the system time of the nth customer, $\forall\ n \geq 1$. Lindley's equation gives the recurrence

$$T_{n+1}(\theta) = \max\{0,\ T_n(\theta) - v_n\} + x_{n+1}(\theta), \quad \forall\ n \geq 1,$$

which governs the evolution of the system. Let $\overline{J}(\theta) = \lim_{n\to\infty} 1/n \sum_{i=1}^{n} J(T_i(\theta),\ \theta)$ be the performance measure, 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. This example has been widely studied in the literature (e.g., [5], [6], [10], [20], [21] [32]–[35]). The problem is to find the optimizer θ^0 such that $\overline{J}(\theta^0) = \min_{\theta\in D} \overline{J}(\theta)$. Noticing that $\overline{J}(\theta)$ does not depend on the initial state of the system, no generality is lost and much convenience is gained by supposing that the first customer arrives at an empty system. In this case, $T_1(\theta) = x_1(\theta)$.

We assume that an analytic formula for $\overline{J}(\theta)$ is not available. A simulation based (or on-line) optimization algorithm, such as the PARMSR algorithm, is then a reasonable option to cope with our problem. We now formulate the PARMSR algorithm updated after the service of every customer. We use the following truncated RM algorithm to update the parameter θ_{n+1}:

$$\begin{cases} \widehat{\theta}_{n+1} = \theta_n - a_n f_{n+1}, \\ \theta_{n+1} = \widehat{\theta}_{n+1} I_{[a<\widehat{\theta}_{n+1}<b]} + aI_{[\widehat{\theta}_{n+1}\leq a]} + bI_{[\widehat{\theta}_{n+1}\geq b]}, \end{cases} \tag{2.1}$$

where f_{n+1} is an estimator of the derivative $f(\theta) \stackrel{\triangle}{=} d\,\overline{J}(\theta)/d\,\theta$ obtained by infinitesimal perturbation analysis (IPA) at the $(n+1)$st step of SA, and $\{a_n\}$ are the step sizes to be chosen. Let Q_n denote the queue length at the instant when the nth customer leaves the server. By the *perturbation propagation rule*, the IPA derivative estimator for $d\,T_{n+1}(\theta)/d\,\theta$ at the $(n+1)$-th step is given by

$$\alpha_{n+1} = \alpha_n\ I_{[Q_n\geq 1]} + \left.\frac{d\,x_{n+1}(\theta)}{d\,\theta}\right|_{\theta=\theta_n}, \quad \alpha_0 = 0, \tag{2.2}$$

where $d\,x_i(\theta)/d\,\theta$, for $i \geq 1$, can be computed by the standard *inversion* method: If $F(\theta,\ x)$ is the distribution function of $x_i(\theta)$ and $\{u_i,\ i \geq 1\}$ is an i.i.d. sequence with a uniform distribution on $(0,1]$, the service time $x_i(\theta)$ is

$$x_i(\theta) = F^{-1}(\theta,\ u_i) = \inf\{x : \quad F(\theta,\ x) \geq u_i\},\ \forall\ i \geq 1,$$

and $d\,x_i(\theta)/d\,\theta$ is the derivative of $x_i(\theta)$ w.r.t. θ for fixed u_i. In the context of on-line optimization, $d\,x_i(\theta)/d\,\theta$ is computed from the observation of $x_i(\theta)$ (for this, we must have partial knowledge of the distribution), and the rest of the PARMSR algorithm works the same. For example, one has $d\,x_i(\theta)/d\,\theta = x_i(\theta)/\theta$ for a scale parameter and $d\,x_i(\theta)/d\,\theta = 1$ for a location parameter. We refer the reader to [12], [16], and [31]–[33] for further details.

Using α_{n+1}, we can compute f_{n+1} as

$$f_{n+1} = J_t(T_{n+1},\ \theta_n)\alpha_{n+1} + J_\theta(T_{n+1},\ \theta_n), \tag{2.3}$$

where T_{n+1} is the system time of the $(n+1)$st customer, while $J_t(\cdot,\ \cdot)$ and $J_\theta(\cdot,\ \cdot)$ denote the partial derivatives of $J(\cdot,\ \cdot)$ with respect to its first and second components, respectively. The observation *noise* (or error) is

$$\varepsilon_{n+1} = f_{n+1} - f(\theta_n). \tag{2.4}$$

The PARMSR algorithm updated after the service of every customer is composed of (2.1)–(2.3).

For our results, let us introduce the following conditions.

(A1). There are three constants $\overline{a} > 0$, $\nu \in (1/2, 1]$, and $\alpha \geq 0$ such that $0 < a_n \leq \overline{a} n^{-\nu}$, $\forall\ n \geq 1$; $a_j = \overline{a},\ \forall\ j \leq 0;\quad \sum_{n=1}^{\infty} a_n = \infty;\quad 0 \leq a_{n+1}^{-1} - a_n^{-1} \xrightarrow[n \to \infty]{} \alpha \geq 0.$

(A2). $f(\theta)$ has a unique root $\theta^0 \in (a,\ b)$, at which $J(\theta)$ is minimized.

(A3). As $\theta \to \theta^0$, $f(\theta)$ can be expressed as $f(\theta) = M_1(\theta - \theta^0) + O(\|\theta - \theta^0\|^2)$ for some constant $M_1 > \alpha/2$, where α is defined in (A1).

(A4). There are constants B_1 to B_5 and μ_0 to μ_2 such that

$$\max\{|J_\theta(t,\ \theta)|,\ |J_t(t,\ \theta)|\} \leq B_1 + B_2 t^{\mu_0},$$

$$\max\{|J_\theta(t_1,\ \theta) - J_\theta(t_2,\ \theta)|,\ |J_t(t_1,\ \theta) - J_t(t_2,\ \theta)|\} \leq B_3|t_1 - t_2| \max(t_1,\ t_2)^{\mu_1},$$

$$\max\{|J_\theta(t,\ \theta_1) - J_\theta(t,\ \theta_2)|,\ |J_t(t,\ \theta_1) - J_t(t,\ \theta_2)|\} \leq B_4|\theta_1 - \theta_2| t^{\mu_2},$$

$$|f(\theta_1) - f(\theta_2)| \leq B_5|\theta_1 - \theta_2|, \quad \forall\ \theta,\ \theta_1,\ \theta_2 \in [a,\ b], \quad t,\ t_1,\ t_2 \in [0,\ \infty).$$

(A5). There exists a measurable function $\Phi(\theta,\ x)$ which does not change sign for $x \geq 0$ and $\theta \in [a,\ b]$, and a polynomial $\sum_{i=0}^{p} b_i x^i, \quad p \geq 1$ such that

$$\frac{d\,x_i(\theta)}{d\,\theta} = \Phi(\theta,\ x_i(\theta)), \quad |\Phi(\theta,\ x)| \leq \sum_{j=0}^{p} b_j x^j, \quad \forall\ \theta \in [a,\ b],\ x \geq 0,$$

$$|\Phi(\theta_1,\ x_i(\theta_1)) - \Phi(\theta_2,\ x_i(\theta_2))| \leq W_0^*|\theta_1 - \theta_2|, \quad \forall\ \theta_1,\ \theta_2 \in [a,\ b],$$

where $\mathrm{E}[W_0^{*4}] < \infty$. Under this condition, $d\,x_i(\theta)/d\,\theta$ does not change sign for $i \geq 1$ and $\theta \in [a,\ b]$. For simplification, we assume that it is positive: $d\,x_i(\theta)/d\,\theta > 0$.

(A6). There are two positive constants β_0 and μ such that $P\{t_1 \leq v_1 \leq t_1 + t\} \leq \beta_0 t^{\mu}$ for all $t, t_1 \geq 0$.

(A7). (i) There exist $p_1 > 1$, $p_2 > 1$, and $q \in [0,\ \mu)$ such that

$$\frac{1}{p_1}\left(1 - \frac{1}{p_2}\right) q > \min\left(\frac{1}{2}, \frac{1}{\nu} - \frac{3}{4}\right), \quad \nu > \frac{2}{3};$$

(ii) $\mathrm{E}[v_1^{\xi_0}] < \infty$, and $\sup_{\theta \in [a,\ b]} \mathrm{E}[x_1(\theta))^{\xi_0}] < \infty$, where

$$\xi_0 = \max\{2(3p + 4\mu_0 + \mu_1 + \mu_2),\ p_2(p + \mu_0),\ \frac{2q}{p_1 - 1}(2p + \mu_0)\}.$$

A typical choice for the sequence $\{a_n\}$ in (A1) is $a_n = \overline{a} n^{-\nu}$, $n \geq 1$. In this case, if $\nu < 1$ then $\alpha = 0$, whereas if $\nu = 1$, then $\alpha = 1/\overline{a}$ and the condition in (A3) becomes $\overline{a} > 1/(2M_1)$, where M_1 represents the second derivative of $\overline{J}(\theta)$ at the optimizer. Conditions (A2) and (A3) imply that $\overline{J}(\theta)$ is unimodal and locally quadratic at θ^0. The polynomial Lipschitz-type assumptions (A4) are mild smoothness conditions on the cost function $J(x,\ \theta)$ and on $f(\theta)$. For the special case where $J(x, \theta) = x + C(\theta)$, one can take $\mu_0 = \mu_1 = \mu_2 = 0$ and $B_2 = B_3 = 0$. The assumptions (A5) on the service time distribution hold, in particular, if θ is a location or scale parameter of the distribution. It is easy to see that $\Phi(\theta, s) \equiv 1$ for a location parameter and $\Phi(\theta, s) = s/\theta$ for a scale parameter (see [33]). For both cases, one can take $W_0^* \equiv 0$ and $p = 1$. Assumption (A6) holds with $\mu = 1$ whenever the interarrival times distribution has a bounded density [35]. (A7) holds whenever the interarrival and service times distributions have bounded moments of large enough order.

For the PARMSR algorithm, let η_{m+1} denote the number of customers served in the $(m+1)$-th BP. We have

$$\eta_{m+1} = \inf\{i :\ \sum_{j=1}^{i} (v_{k_m+j} - x_{k_m+j}(\theta_{k_m+j-1})) > 0\}, \tag{2.5}$$

where $k_0 = 0$ and $k_m \triangleq \sum_{i=1}^{m} \eta_i$ for $m \geq 1$. Let C_i denote the ith customer entering the system. In the $(m+1)$-th BP, if θ_{k_m+1}, θ_{k_m+2}, ... are replaced by θ, we denote the number of customers in the $(m+1)$-th BP by $\eta_{m+1}(\theta)$. In particular, if the service parameters are fixed at b throughout the $(m+1)$-th BP, then there may be some additional customers served in this BP after customer $C_{k_{m+1}}$. Denote them by $C^{(1)}_{k_{m+1}+1}$, ..., $C^{(1)}_{k_m+\eta_{m+1}(b)}$. The corresponding additional random numbers needed are $v^{(1)}_{k_{m+1}+1}$, ..., $v^{(1)}_{k_m+\eta_{m+1}(b)}$; $u^{(1)}_{k_{m+1}+1}$, ..., $u^{(1)}_{k_m+\eta_{m+1}(b)}$. For simplicity of writing, they are still denoted by $v_{k_{m+1}+1}$, ..., $v_{k_m+\eta_{m+1}(b)}$; $u_{k_{m+1}+1}$, ..., $u_{k_m+\eta_{m+1}(b)}$. Let the length of the $(m+1)$-th busy period (BP) and the $(m+1)$-th idle period be L_{m+1} and I_{m+1}, respectively, if the PARMSR algorithm is applied, and let these lengths be $L_{m+1}(\theta)$ and $I_{m+1}(\theta)$ if the service parameters are fixed at θ throughout the $(m+1)$-th BP. More precisely,

$$L_{m+1}(\theta)=\sum_{i=1}^{\eta_{m+1}(\theta)} x_{k_m+i}(\theta), \qquad L_{m+1}=\sum_{i=1}^{\eta_{m+1}} x_{k_m+i}(\theta_{k_m+i-1}),$$

$$I_{m+1}(\theta)=\sum_{i=1}^{\eta_{m+1}(\theta)} (v_{k_m+i}-x_{k_m+i}(\theta)), \quad I_{m+1}=\sum_{i=1}^{\eta_{m+1}} (v_{k_m+i}-x_{k_m+i}(\theta_{k_m+i-1})).$$

We introduce the following additional condition.

(A8). $\overline{\eta}(\theta)\stackrel{\triangle}{=}\mathrm{E}[\eta_1(\theta)]$ and $\mathrm{E}\left[\left(\sum_{j=1}^{\eta_1(\theta)} f_j(\theta)\right)^2\right]$ are continuous at the optimal point θ^0, where $f_j(\theta)=J_t(T_j(\theta),\ \theta)\sum_{l=1}^{j}\frac{d\,x_l(\theta)}{d\,\theta}+J_\theta(T_j(\theta),\ \theta)$ for $\theta\in[a,\ b]$ and $1\le j\le\eta_1(\theta)$.

L'Ecuyer and Glynn [20], in their Proposition 15, show how (A8) can be verified from simple conditions on the service time distribution.

The main result in this section is the following theorem, whose proof is given in the appendix. Define $A=-M_1+\alpha/2$ (from (A3), one has $A<0$),

$$\sigma^2=\mathrm{E}\left[\left(\sum_{j=1}^{\eta_1(\theta^0)} f_j(\theta^0)\right)^2\right], \qquad \text{and} \quad \sigma_1^2=\frac{\sigma^2}{\overline{\eta}(\theta^0)}\int_0^\infty e^{2At}dt=\frac{\sigma^2}{(2M_1-\alpha)\overline{\eta}(\theta^0)}.$$

Theorem 2.1 *If conditions (A1)–(A8) hold, then the PARMSR algorithm defined by (2.1)–(2.3) satisfies*

$$\frac{\theta_n-\theta^0}{\sqrt{a_n}}\xrightarrow[n\to\infty]{d}\mathcal{N}(0,\ \sigma_1^2),$$

where $\mathcal{N}(0,\ \sigma_1^2)$ means a normal distribution with mean 0 and variance σ_1^2.

3 Central Limit Theorem for the PARMSR Algorithm Updated After Every L Customers

We now turn to the case in which the PA derivative estimator at each step of the SA algorithm is based on L successive customers. When $L=1$, this reduces to the setup of Section 2. Our intention is to explore how the limiting behavior of the PARMSR algorithm depends on the choice of step sizes, the value of L, and the average number of customers per BP.

At the $(n+1)$-th step, the derivative estimator is defined by

$$f_{n+1}=\frac{1}{L}\sum_{i=1}^{L}[J_t(T_{nL+i},\ \theta_n)\beta_{n+1,i}+J_\theta(T_{nL+i},\ \theta_n)], \tag{3.1}$$

where $\beta_{n+1,i}$, $1 \le i \le L$ are the estimators for the sample path derivatives of the system times T_{nL+i}, $1 \le i \le L$. By the *perturbation propagation rule*, $\beta_{n+1,i}$, $1 \le i \le L$ are defined by

$$\begin{cases} \beta_{n+1,0} = \beta_{n,L}\ , \quad \beta_{1,0} = 0, \\ \beta_{n+1,i+1} = \beta_{n+1,i}\ I_{[Q_{nL+i} \ge 1]} + \left.\dfrac{d\,x_{nL+i+1}(\theta)}{d\,\theta}\right|_{\theta=\theta_n}, \quad 0 \le i \le L-1. \end{cases} \tag{3.2}$$

The PARMSR algorithm updated after every L-customers is defined by (2.1), (3.1) and (3.2). This algorithm has been widely investigated; see, e.g., [20], [21], [32]–[35].

Theorem 3.1 *Suppose that conditions (A1)–(A8) are fulfilled. Then for the algorithm defined by (2.1), (3.1) and (3.2) one has*

$$\frac{\theta_n - \theta^0}{\sqrt{a_n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\ \sigma_2^2),$$

where $\sigma_2^2 = \sigma_1^2/L = \sigma^2/(L\overline{\eta}(\theta^0)(2M_1 - \alpha))$.

The proof of Theorem 3.1 is in the Appendix. We want to express the convergence in terms of the total number of customers (the computing budget), which is here $N_n = nL$ after n SA iterations. For this, we assume that $a_n = \overline{a}n^{-\nu}$ for all n, where $\nu \in (1/2,\ 1]$. Then it is seen that

$$a_n\sigma_2^2 = \frac{\overline{a}\sigma^2 n^{-\nu}}{(2M_1 - \alpha)\overline{\eta}(\theta^0)L} = \frac{\overline{a}\sigma^2}{L^{1-\nu}(2M_1 - \alpha)\overline{\eta}(\theta^0)} N_n^{-\nu}. \tag{3.3}$$

The result of Theorem 3.1 can then be rewritten, in terms of the criterion (1.2) discussed in the introduction, as

$$\frac{N_n^{\nu/2}(\theta_n - \theta^0)}{\kappa_1} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1), \quad \text{where } \kappa_1^2 = \sigma^2 \frac{\overline{a}}{L^{1-\nu}(2M_1 - \alpha)\overline{\eta}(\theta^0)}. \tag{3.4}$$

The convergence rate is optimized by taking $\nu = 1$, and then the variance constant $\overline{a}\sigma^2/((2M_1 - \alpha)\overline{\eta}(\theta^0))$ is independent of L and has a minimum value of $\sigma^2/(M_1^2\overline{\eta}(\theta^0))$, attained by choosing $\overline{a} = 1/M_1$. For $\nu < 1$, the limiting variance constant decreases with L. Of course, this tells only the asymptotic story. For a finite N_n, a large L can be bad because it allows too few SA iterations and thus keeps the variance large. This is summarized by:

Corollary 3.1 *Under conditions (A1)–(A8), the optimal choice for the step-sizes* $\{a_n,\ n \ge 1\}$ *in the PARMSR algorithm updated after every* L *customers is* $a_n = 1/(nM_1)$. *With this choice, one has*

$$\frac{M_1\sqrt{\overline{\eta}(\theta^0)N_n}(\theta_n - \theta^0)}{\sigma} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1). \tag{3.5}$$

Taking $\nu < 1$ *deteriorates the convergence rate, while taking* $\overline{a} \ne 1/M_1$ *with* $\nu = 1$ *deteriorates the variance constant if* $\overline{a} > 1/(2M_1)$ *and can also change the convergence rate if* $\overline{a} \le 1/(2M_1)$.

The derivative estimator (3.1) is an average over L customers and its variance is roughly proportional to $1/L$ (at least for large L). So, spending more time at each SA iteration decreases the variance proportionally, and this explains why the variance constant in (3.5), where things are expressed in terms of the total computing time N_n, is independent of L.

What if we replace the $1/L$ in front of the sum in (3.1) by an arbitrary constant κ? In lieu of (3.1), define the $(n+1)$-th *rescaled* estimator of $f(\theta)$ by

$$f_{n+1} = \kappa \sum_{i=1}^{L} [J_t(T_{nL+i},\ \theta_n)\beta_{n+1,i} + J_\theta(T_{nL+i},\ \theta_n)], \quad \forall n \geq 0. \tag{3.6}$$

Theorem 3.2 *Suppose that the conditions (A1)–(A8) are fulfilled, with the condition* $-M_1+\alpha/2 < 0$ *replaced by* $-\kappa L M_1 + \alpha/2 < 0$*. Then for the modified PARMSR algorithm updated after every* L *customers, defined by (2.1), (3.6) and (3.2), one has*

$$\frac{\theta_n - \theta^0}{\sqrt{a_n}} \xrightarrow[n \to \infty]{d} \mathcal{N}(0,\ \sigma_3^2),$$

where

$$\sigma_3^2 = \kappa L \sigma_2^2 \left(\frac{-M_1 + \alpha/2}{-M_1 + \alpha/(2\kappa L)} \right) = \frac{\kappa^2 L \sigma^2}{\overline{\eta}(\theta^0)(2M_1\kappa L - \alpha)}.$$

Proof. Since $a_n f_{n+1} = (\kappa L a_n)(f_{n+1}/(\kappa L))$, we can view the modified PARMSR algorithm updated after every L customers as a special case of the original algorithm with the step-sizes $\{a_n\}$ replaced by $\{\kappa L a_n\}$. Thus Theorem 3.2 follows easily from Theorem 3.1 if we replace $\{a_n\}$ and α in Theorem 3.1 by $\{\kappa L a_n\}$ and $\alpha/(\kappa L)$ respectively.

□

If we take $a_n = \overline{a} n^{-1}$ in this setup, then

$$a_n \sigma_3^2 = \frac{\overline{a}\kappa^2 L \sigma^2}{n(2M_1\kappa L - \alpha)\overline{\eta}(\theta^0)} = \frac{\overline{a}\kappa^2 L^2 \sigma^2}{N_n(2M_1\kappa L - \alpha)\overline{\eta}(\theta^0)}.$$

It has a minimum value of $\sigma^2/(M_1^2\overline{\eta}(\theta^0))$, attained by choosing $\overline{a} = 1/(M_1\kappa L)$. With this choice, the central limit theorem (3.5) holds, just as in Corollary 3.1, independently of κ and L. We note that rescaling the function f or its estimator by a constant is equivalent to rescaling a_n by the same constant. In other words, if we replace (3.1) by (3.6), the effect is the same as multiplying a_n by κL, so it suffices to analyze the result of Theorem 3.1 with this change of a_n to understand what would happen. This scaling issue is well-known in stochastic approximation.

4 Comparison with the PARMSR Algorithm Updated After Every Busy Period

We first formulate the procedure for the PARMSR algorithm updated after every busy period. Since $\overline{\eta}(\theta) \geq 1$ for all $\theta \in [a,\ b]$, finding a root of $f(\theta)$ is equivalent to finding a root of $f(\theta)\overline{\eta}(\theta)$. At the

$(n+1)$-th step of the SA algorithm, we use θ_n as the service parameter throughout the $(n+1)$-th BP. At the end of the BP, we compute θ_{n+1} based on the evolution of the whole BP. The $(n+1)$-th estimator for $f(\theta)\overline{\eta}(\theta)$ is defined by

$$f_{n+1} = \sum_{i=1}^{\eta_{n+1}(\theta_n)} f_{k_n+i}(\theta_n), \tag{4.1}$$

where $f_{k_n+i}(\theta)$ is given by (6.22) of the appendix and $k_n = \sum_{i=1}^{n} \eta_i(\theta_{i-1})$.

The PARMSR algorithm updated after every BP is composed of (2.1) and (4.1); see, e.g., [5], [10], [20], and [21].

Theorem 4.1 *Suppose that conditions (A1)–(A5) and (A7)–(A8) are satisfied with the condition* $-M_1+\alpha/2 < 0$ *replaced by* $-M_1\overline{\eta}(\theta^0)+\alpha/2 < 0$ *and the condition on* ξ_0 *replaced by* $\xi_0 > 2(p+\mu_0+1)$. *Then the algorithm defined by (2.1) and (4.1) satisfies*

$$\frac{\theta_n - \theta^0}{\sqrt{a_n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\ \sigma_4^2),$$

where $\sigma_4^2 = \sigma^2/(2M_1\overline{\eta}(\theta^0) - \alpha)$.

Proof. Let $\varepsilon_{n+1} = f_{n+1} - \overline{\eta}(\theta_n)f(\theta_n)$. It is seen that $\{\varepsilon_n,\ \mathcal{F}_{k_n}\}$ is a sequence of martingale differences. The rest the proof is easily derived from (6.49) and is standard (see, e.g., [4, 17, 24]).

□

We now express the convergence in terms of the computing budget for n SA iterations, which is a random variable $N_n = k_n = \sum_{i=1}^{n} \eta_i(\theta_{i-1})$. Since $\mathrm{E}[\eta_1(\theta_n)] \to \overline{\eta}(\theta^0)$ as $\theta_n \to \theta^0$ by (A8), using a standard martingale argument it can be shown that the computing budget N_n for n SA iterations obeys $N_n/n \to \overline{\eta}(\theta^0)$ a.s. as $n \to \infty$. Taking $a_n = \overline{a}n^{-\nu}$, $\forall\ n \geq 1$ where $\nu \in (1/2,\ 1]$, one has

$$\sigma_4^2 a_n = \frac{\overline{a}\sigma^2 n^{-\nu}}{2M_1\overline{\eta}(\theta^0) - \alpha} \sim \frac{\overline{a}\sigma^2\overline{\eta}(\theta^0)^{\nu}}{2M_1\overline{\eta}(\theta^0) - \alpha} N_n^{-\nu}, \quad a.s., \tag{4.2}$$

where the notation $y_n^{(1)} \sim y_n^{(2)}$ means that $y_n^{(1)}/y_n^{(2)} \to 1$ as $n \to \infty$. By (4.2), the result in Theorem 3.1 can also be expressed in the form of the criterion (1.2) as follows:

$$\frac{N_n^{\nu/2}(\theta_n - \theta^0)}{\kappa_2} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1), \quad \text{where } \kappa_2^2 = \sigma^2 \frac{\overline{a}\overline{\eta}^{\nu}(\theta^0)}{2M_1\overline{\eta}(\theta^0) - \alpha}. \tag{4.3}$$

Again, $\nu = 1$ gives the best convergence rate, and $\overline{a} = 1/(M_1\overline{\eta}(\theta^0))$ minimizes the limit variance constant κ_2^2, with minimal value equal to $\sigma^2/(M_1^2\overline{\eta}(\theta^0))$. We thus have

Corollary 4.1 *Under the conditions of Theorem 4.1, the optimal choice for the sequence* $\{a_n\}$ *in the PARMSR algorithm updated after every BP is* $a_n = 1/(nM_1\overline{\eta}(\theta^0))$. *With this choice, one has*

$$\frac{M_1\sqrt{\overline{\eta}(\theta^0)N_n}(\theta_n - \theta^0)}{\sigma} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1),$$

which is exactly the same as for the PARMSR algorithm updated after every L customers. The convergence rate deteriorates with $\nu < 1$, whereas the limiting variance constant deteriorates if $\nu = 1$ but $\overline{a} > 1/(2M_1\overline{\eta}(\theta^0))$ and $\overline{a} \neq 1/(M_1\overline{\eta}(\theta^0))$.

We see that when the convergence rate is expressed in terms of N_n, which is the appropriate thing to do, the algorithm updated after every L customers and the algorithm updated after every BP both have the same convergence rate and the same variance constant, provided that the sequence $\{a_n,\ n \geq 1\}$ is chosen optimally in each case. But their optimal step sizes are different. This agrees with the numerical results given in [21].

If both algorithms use the same step sizes a_n, of course they have a different limiting behavior, and the algorithm that does best in this situation is the one whose step sizes $\{a_n\}$ match best the optimal step sizes. Let $a_n = \overline{a}n^{-\nu}$, $\forall\ n \geq 1$ where $\nu \in (1/2,\ 1]$ is a constant. By (3.4) and (4.3), it is easy to see that

$$\kappa_2 > \kappa_1 \quad \Longleftrightarrow \quad \overline{\eta}^{\nu}(\theta^0)(2M_1 - \alpha)L^{1-\nu} > 2M_1 - \alpha/\overline{\eta}(\theta^0), \tag{4.4}$$

assuming that the two expressions on the right side are positive. From (4.4) we find that if α is small then $\kappa_2 > \kappa_1$, whereas if α is near $2M_1$, then $\kappa_2 < \kappa_1$. Notice that if we choose $\nu < 1$, then $\alpha = 0$. In this case, it is obvious that $\kappa_2 > \kappa_1$. If $\alpha \geq 2M_1$, then our results on the PARMSR algorithm updated after every L customers are not applicable. In this case, however, Theorem 4.1 may still hold since $2M_1\overline{\eta}(\theta^0) > \alpha$ may be satisfied. In general, one can use an algorithm updated after every L^* BPs if the condition $L^*M_1\overline{\eta}(\theta^0) > \alpha$ is satisfied.

We note that weaker hypotheses are imposed in Theorem 4.1 than that in Theorem 2.1. The PARMSR algorithm updated after every BP in this section can in principle be extended to regenerative systems in a wide sense (i.e., Harris-recurrent models; see, e.g., Asmussen [1]), if explicit regeneration points can be found before the implementation of the algorithm. We do not pursue this generality in this paper.

5 Conclusion

Central limit theorems have been established for the PARMSR algorithms updated either after every L customers or after every BP, for the $GI/GI/1$ queueing systems. It has been shown how the limiting behavior of the two kinds of algorithms depends on the choice of step sizes, the value of the observation period between updates, and the average number of customers in a BP. Our proofs are given only for this special model but they could serve as a basis for more general proofs in the context of controlled regenerative processes.

Our results show that the step sizes of the algorithms play a major role in their limiting behavior. The idea of averaging the iterates in stochastic approximation (see, e.g., [2], [4], [26], [38]), which is devoted to achieving an optimal convergence rate and optimal variance constant for a broader range of step sizes, can also be combined in a natural way with the PARMSR algorithms. Analyzing its convergence properties in terms of the computing budget, in the context of infinite-horizon models, is an ongoing topic of research.

When implementing the PARMSR algorithm updated after every fixed-length observation period for a more complicated DEDS, such as a queueing network with many nodes, an important concern is the transient effects introduced in the system by frequently changing the parameters, as explained in [31]–[33]. If the controlled parameters are fixed at a constant, a queueing system may enter the steady state at finite time (see, e.g., [1]). However, with a changing parameter, the system is no longer stationary. In particular, a regenerative system does not necessarily remain regenerative in the case where the control parameters are changed frequently. The transient effects may also induce a change in the order of events, compared with the case where the control parameters are fixed, because the parameter change is *not* infinitesimal (in contrast to PA, where the order of events remains the same because the parameter change is infinitesimal; see, e.g., [12, 16, 19]). In order to broaden the range of application of the PARMSR algorithm updated after every fixed-length observation period, the above issues need to be properly addressed. Although the findings of this paper as well as other results have improved our understanding of convergence of the PARMSR algorithms, further research focusing on parameter optimization for more complex systems, such as the closed and the open queueing networks with many nodes, is still needed.

6 Appendix

This appendix contains the proofs of the theorems.

In order to prove Theorem 2.1, we first need several lemmas. Set

$$\sigma(n) = \sup\{j : \quad k_j \le n\}, \quad \tau(n) = \sigma(n) + 1 = \inf\{j : \quad k_j > n\}; \tag{6.1}$$

$$\left.\begin{array}{l} \varphi_{n,\,k} = \prod_{i=k}^{n}(1 + a_i A_i), \quad \forall\ n \ge k; \\ \varphi_{0,0} = 1,\ \varphi_{n,\ n+1} = 1; \quad \varphi_{i,j} = 0,\ \forall\ j \ge i+2, \end{array}\right\} \tag{6.2}$$

where $\{A_n,\ n \ge 1\}$ is a deterministic sequence and satisfies $\lim_{n\to\infty} A_n = A = -M_1 + \alpha/2 < 0$.

Lemma 6.1 *If conditions (A1)–(A8) are fulfilled, then*

$$\sum_{j=0}^{\sigma(n)} a_{k_j} \varphi^2_{n,\,k_j} \xrightarrow[n\,\to\,\infty]{P} \frac{1}{\overline{\eta}(\theta^0)} \int_0^\infty e^{2At} dt,$$

where $\xrightarrow{P}$ *means convergence in probability.*

Proof. Since $\lim_{n\to\infty} A_n = A < 0$, it is standard to derive that

$$|\varphi_{n,\ k}| \le c_0 \exp\{-c \sum_{j=k}^{n} a_j\}, \tag{6.3}$$

$$\sup_n \sum_{i=1}^n a_i |\varphi_{n,\ i}|^r < \infty, \quad \forall\ r > 0, \tag{6.4}$$

$$\lim_{n\to\infty} \sum_{i=0}^n a_i \varphi_{n,\ i+1}^2 = \int_0^\infty e^{2At} dt, \tag{6.5}$$

where c_0 and c are some positive constants (see, e.g., [2, 4, 24, 38]). By the definition (6.2), one has

$$\sum_{i=0}^n a_i \varphi_{n,\ i+1}^2 = \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1} \varphi_{n,\ k_m+i}^2,$$

which, together with (6.5), gives

$$\lim_{n\to\infty} \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1} \varphi_{n,\ k_m+i}^2 = \int_0^\infty e^{2At} dt. \tag{6.6}$$

We now decompose

$$\begin{aligned}
& \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1} \varphi_{n,\ k_m+i}^2 \\
= \ & \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} (a_{k_m+i-1} - a_{k_m}) \varphi_{n,\ k_m+i}^2 + \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m} (\varphi_{n,\ k_m+i}^2 - \varphi_{n,\ k_m}^2) \\
& + \sum_{m=0}^{\sigma(n)} a_{k_m} \varphi_{n,\ k_m}^2 (\eta_{m+1} - \eta_{m+1}(\theta_{k_m})) + \sum_{m=0}^{\sigma(n)} a_{k_m} \varphi_{n,\ k_m}^2 (\eta_{m+1}(\theta_{k_m}) - \overline{\eta}(\theta_{k_m})) \\
& + \sum_{m=0}^{\sigma(n)} a_{k_m} \varphi_{n,\ k_m}^2 (\overline{\eta}(\theta_{k_m}) - \overline{\eta}(\theta^0)) \ + \ \overline{\eta}(\theta^0) \sum_{m=0}^{\sigma(n)} a_{k_m} \varphi_{n,\ k_m}^2 .
\end{aligned} \tag{6.7}$$

In the sequel, we will prove that each of the first five terms on the right hand side of (6.7) tends to 0 in some sense.

(i). By (A1), there is a constant α_0^* such that

$$0 \le a_n^{1-\delta'} - a_{n+1}^{1-\delta'} \le \alpha_0^* a_n^{2-\delta'}, \quad \forall\ n \ge 1,\ \delta' \in [0, 1),$$

which gives

$$\max_{1\le i\le \eta_{m+1}} |a_{k_m+i-1}^{1-\delta'} - a_{k_m}^{1-\delta'}| \le \alpha_0^* \eta_{m+1}(b) a_{k_m}^{2-\delta'}, \quad \forall\ \delta' \in [0, 1). \tag{6.8}$$

By condition (A7) (ii), we have

$$\mathrm{E}\left[\eta_1(b)^{\xi_0}\right] < \infty, \quad \mathrm{E}\left[L_1(b)^{\xi_0}\right] < \infty; \tag{6.9}$$

see, e.g., [13] and [36]. By (A1), (6.9), and Lemma 2 of Tang and Chen [34] it follows that

$$a_{k_m} \eta_{m+1}(b) \xrightarrow[m\ \to\ \infty]{} 0, \quad a.s. \tag{6.10}$$

Since $\sum_{n=1}^{\infty} a_n = \infty$, from (6.3) it follows that $|\varphi_{n,\ k'}| \xrightarrow[n\to\infty]{} 0$ for any fixed k'. By (6.6), (6.8), and (6.10),

$$
\begin{aligned}
&|\sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} (a_{k_m+i-1} - a_{k_m})\varphi^2_{n,\ k_m+i}| \le \alpha_0^* \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a^2_{k_m} \varphi^2_{n,\ k_m+i}\eta_{m+1}(b) \\
\le\ & \alpha_0^* \sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1}\varphi^2_{n,\ k_m+i}(\max_{1\le i\le \eta_{m+1}} \frac{a_{k_m}}{a_{k_m+i-1}}) a_{k_m}\eta_{m+1}(b) \\
& \xrightarrow[n\to\infty]{} 0, \quad a.s.
\end{aligned}
\tag{6.11}
$$

(ii). From the definition (6.2) we have

$$\varphi_{n,\ k+1} - \varphi_{n,\ k} = \varphi_{n,\ k+1}(-a_k A_k), \quad \forall\ n \ge k,$$

which yields

$$
\begin{aligned}
&|\varphi^2_{n,\ k_m+i} - \varphi^2_{n,\ k_m}| = |\sum_{j=1}^{i}(\varphi_{n,\ k_m+j} - \varphi_{n,\ k_m+j-1})(\varphi_{n,\ k_m+i} + \varphi_{n,\ k_m})| \\
\le\ & |\varphi_{n,\ k_m+i} + \varphi_{n,\ k_m}| \sum_{j=1}^{i} a_{k_m+j-1}|\varphi_{n,\ k_m+j}A_{k_m+j-1}| \\
\le\ & \alpha_1^* \varphi^2_{n,\ k_m+i} a_{k_m}\eta_{m+1}(b), \quad \forall\ m \ge m_0,\ 1 \le i \le \eta_{m+1},
\end{aligned}
\tag{6.12}
$$

for some constant α_1^* and sufficiently large m_0.

By (6.3), (6.6), (6.8) and (6.12), analogous to (6.11), it is derived that

$$
\begin{aligned}
&|\sum_{m=0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m}(\varphi^2_{n,\ k_m+i} - \varphi^2_{n,\ k_m})| \\
\le\ & |\sum_{m=0}^{m_0-1} \sum_{i=1}^{\eta_{m+1}} a_{k_m}(\varphi^2_{n,\ k_m+i} - \varphi^2_{n,\ k_m})| + \alpha_1^* \sum_{m=m_0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a^2_{k_m}\varphi^2_{n,\ k_m+i}\eta_{m+1}(b) \\
=\ & |\sum_{m=0}^{m_0-1} \sum_{i=1}^{\eta_{m+1}} a_{k_m}(\varphi^2_{n,\ k_m+i} - \varphi^2_{n,\ k_m})| \\
& + \alpha_1^* \sum_{m=m_0}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1}\varphi^2_{n,\ k_m+i}(\max_{1\le i\le \eta_{m+1}} \frac{a_{k_m}}{a_{k_m+i-1}}) a_{k_m}\eta_{m+1}(b) \xrightarrow[n\to\infty]{} 0, \quad a.s.
\end{aligned}
\tag{6.13}
$$

(iii). By (6.10), in order to prove

$$\sum_{m=0}^{\sigma(n)} a_{k_m}\varphi^2_{n,\ k_m}(\eta_{m+1} - \eta_{m+1}(\theta_{k_m})) \xrightarrow[n\to\infty]{P} 0, \tag{6.14}$$

it suffices to prove

$$\sum_{m=0}^{\tau(n)} a_{k_m}\varphi^2_{n,\ k_m}(\eta_{m+1} - \eta_{m+1}(\theta_{k_m})) \xrightarrow[n\to\infty]{P} 0. \tag{6.15}$$

By Lemma 3.3 of [35], there exist constant $\beta_2 > 0$ and $\gamma_1 \in (0,\ \mu)$, where μ is given in condition (A6), such that

$$P\{\eta_{m+1} \neq \eta_{m+1}(\theta_{k_m})|\mathcal{F}_{k_m}\} \leq \beta_2 a_{k_m}^{\gamma_1}, \quad \forall\ m \geq 1, \tag{6.16}$$

where the filtration $\mathcal{F}_{k_m}$ is defined by

$$\mathcal{F}_{k_m} = \sigma\{B:\ B \in \sigma\{v_j,\ u_j,\ \forall\ j \geq 1\},\ B[k_m = i] \in \sigma\{v_1,\ u_1;\ ...;\ v_i,\ u_i\},\ \forall\ i \geq 1\}.$$

By (6.16) and the Schwarz inequality it is derived that

$$\begin{aligned}
& \mathrm{E}\left[\left|\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2(\eta_{m+1} - \eta_{m+1}(\theta_{k_m}))\right|\right] \\
\leq\ & \mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2\eta_{m+1}(b)I\{\eta_{m+1} \neq \eta_{m+1}(\theta_{k_m})\}\right] \\
=\ & \mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2\mathrm{E}\{\eta_{m+1}(b)I\{\eta_{m+1} \neq \eta_{m+1}(\theta_{k_m})\}|\mathcal{F}_{k_m}\}\right] \\
\leq\ & \mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2\sqrt{\mathrm{E}[\eta_{m+1}(b)^2]}\sqrt{\beta_2}a_{k_m}^{\gamma_1/2}\right] \\
\leq\ & \sqrt{\beta_2}\sqrt{\mathrm{E}[\eta_1(b)^2]}\mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2\overline{a}^{\gamma_1/2}m^{-\nu\gamma_1/2}\right].
\end{aligned} \tag{6.17}$$

Since $\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2 \leq \sum_{m=0}^{\tau(n)} a_{k_m-1}\varphi_{n,\ k_m}^2 \leq \sum_{i=-1}^{n} a_i\varphi_{n,\ i+1}^2$, by (6.5) we conclude that $\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2$, $\forall\ n \geq 1$ are bounded above by some constant. By (6.3), it follows from (6.17) that

$$\lim_{n\to\infty} \mathrm{E}\left[\left|\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2(\eta_{m+1} - \eta_{m+1}(\theta_{k_m}))\right|\right] = 0, \tag{6.18}$$

which implies (6.15).

(iv). Notice that $\{\eta_{m+1}(\theta_{k_m}) - \overline{\eta}(\theta_{k_m}),\ \mathcal{F}_{k_m}\}$ is a sequence of martingale differences and that

$$\mathrm{E}[(\eta_{m+1}(\theta_{k_m}) - \overline{\eta}(\theta_{k_m}))^2|\mathcal{F}_{k_m}] \leq \mathrm{E}[\eta_1(b)^2] < \infty.$$

By Burkholder's inequality (see, e.g., [9] and [14]), there exists a constant B_1^* such that

$$\begin{aligned}
& \mathrm{E}\left[\left(\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^2(\eta_{m+1}(\theta_{k_m}) - \overline{\eta}(\theta_{k_m}))\right)^2\right] \\
\leq\ & B_1^*\mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}^2\varphi_{n,\ k_m}^4\mathrm{E}[(\eta_{m+1}(\theta_{k_m}) - \overline{\eta}(\theta_{k_m}))^2|\mathcal{F}_{k_m}]\right] \\
\leq\ & B_1^*\mathrm{E}[\eta_1(b)^2]\mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}\varphi_{n,\ k_m}^4\overline{a}m^{-\nu}\right] \xrightarrow[n\ \to\ \infty]{} 0,
\end{aligned} \tag{6.19}$$

via the same argument as in (6.18). From (6.19) we have

$$\sum_{m=0}^{\sigma(n)} a_{k_m}\varphi^2_{n,\ k_m}(\overline{\eta}(\theta_{k_m})-\overline{\eta}(\theta_{k_m}))\xrightarrow[n\to\infty]{P}0. \tag{6.20}$$

(v). Under conditions (A1)–(A7), by Theorem 2.1 of [35], $\lim\limits_{n\to\infty}\theta_n=\theta^0$ a.s. Since $\overline{\eta}(\theta)$ is continuous at θ^0 by condition (A8), it is seen that $\lim\limits_{m\to\infty}\overline{\eta}(\theta_{k_m})=\overline{\eta}(\theta^0)$, a.s. By the boundedness of $\sum\limits_{m=0}^{\sigma(n)} a_{k_m}\varphi^2_{n,\ k_m}$, it is readily seen that

$$\sum_{m=0}^{\sigma(n)} a_{k_m}\varphi^2_{n,\ k_m}(\overline{\eta}(\theta_{k_m})-\overline{\eta}(\theta^0))\xrightarrow[n\to\infty]{}0,\quad a.s. \tag{6.21}$$

Substituting (6.11), (6.13), (6.14), (6.20), and (6.21) into (6.7), by (6.6) the lemma follows.

□

Define

$$\left.\begin{aligned}
&f_{k_m+i}(\theta)=J_t(T_{k_m+i}(\theta),\ \theta)\alpha_{k_m+i}(\theta)+J_\theta(T_{k_m+i}(\theta),\ \theta),\\
&T_{k_m+i}(\theta)=\sum_{j=1}^{i}x_{k_m+j}(\theta)-\sum_{j=1}^{i-1}v_{k_m+j},\\
&\alpha_{k_m+i}(\theta)=\sum_{j=1}^{i}\frac{d\,x_{k_m+j}(\theta)}{d\,\theta},\quad \forall\ \theta\in[a,\ b],\ 1\le i\le\eta_{m+1}(\theta),\ m\ge 0.
\end{aligned}\right\} \tag{6.22}$$

In the above definitions, $\alpha_{k_m+i}(\theta)$ is the derivative of the system time $T_{k_m+i}(\theta)$ of the (k_m+i)-th customer, if the service parameters are fixed at θ throughout the $(m+1)$-th BP, and $f_{k_m+i}(\theta)$ is an estimator of $f(\theta)$.

Lemma 6.2 *Let* $w_{m+1}=\sum\limits_{i=1}^{\eta_{m+1}(\theta_{k_m})}(f_{k_m+i}(\theta_{k_m})-f(\theta_{k_m})),\ \forall\ m\ge 0.$ *If conditions (A1)–(A8) are satisfied, then for the PARMSR algorithm updated after the service of every customer defined by (2.1)–(2.3) the following result holds*

$$\sum_{m=0}^{\sigma(n)}\sqrt{a_{k_m}}\varphi_{n,\ k_m}w_{m+1}\xrightarrow[n\to\infty]{d}\mathcal{N}(0,\ \sigma_1^2). \tag{6.23}$$

Proof. As in Lemma 3.1 of [35], it can be shown that

$$\sqrt{a_{k_m}}w_{m+1}\xrightarrow[n\to\infty]{}0,\quad a.s.,$$

which gives

$$\sqrt{a_{k_{\sigma(n)+1}}}\varphi_{n,\ k_{\sigma(n)+1}}w_{\sigma(n)+2}\xrightarrow[n\to\infty]{}0,\quad a.s.$$

By the definition (6.2), it is seen that

$$\begin{aligned}
&|\sum_{m=0}^{n}\sqrt{a_{k_m}}\varphi_{n,\ k_m}w_{m+1}-\sum_{m=0}^{\sigma(n)}\sqrt{a_{k_m}}\varphi_{n,\ k_m}w_{m+1}|\\
\le\ &\sqrt{a_{k_{\sigma(n)+1}}}|\varphi_{n,\ k_{\sigma(n)+1}}w_{\sigma(n)+2}|\xrightarrow[n\to\infty]{}0,\quad a.s.
\end{aligned}$$

This means that in order to prove (6.23), it suffices to prove

$$\sum_{m=0}^{n} \sqrt{a_{k_m}} \varphi_{n,\ k_m} w_{m+1} \xrightarrow[n \to \infty]{d} \mathcal{N}(0,\ \sigma_1^2). \tag{6.24}$$

Let $X_{n,\ j} = \sqrt{a_{k_j}} \varphi_{n,\ k_j} w_{j+1}, \quad 1 \le j \le n$. We will verify that the conditions of Theorem 9.5.2 in Chow and Teicher [9] hold for $\sum_{j=1}^{n} X_{n,\ j}$.

As in the proof of Lemma 3.4 (part (ii)) in [35], it is shown that $\{w_{j+1},\ \mathcal{F}_{k_j},\ j \ge 0\}$ is a sequence of martingale differences. Thus we have

$$\mathrm{E}[X_{n,\ j} | \mathcal{F}_{k_j}] = 0, \quad 1 \le j \le n,$$

which yields

$$\sum_{j=1}^{n} \mathrm{E}[X_{n,j} | \mathcal{F}_{k_j}] = 0, \quad a.s. \tag{6.25}$$

By condition (A8), $\Delta(\theta) = \mathrm{E}\left[\left(\sum_{i=1}^{\eta_1(\theta)} (f_i(\theta) - f(\theta))\right)^2\right]$ is continuous at θ^0. Since $\theta_{k_m} \xrightarrow[m \to \infty]{} \theta^0$ a.s., it follows that

$$\mathrm{E}[w_{j+1}^2 | \mathcal{F}_{k_j}] = \Delta(\theta_{k_j}) \xrightarrow[j \to \infty]{} \sigma^2 = \mathrm{E}\left[\left(\sum_{i=1}^{\eta_1(\theta^0)} f_i(\theta^0)\right)^2\right],$$

which yields

$$\sum_{j=0}^{n} \mathrm{E}[X_{n,\ j}^2 | \mathcal{F}_{k_j}] = \sum_{j=0}^{n} a_{k_j} \varphi_{n,\ k_j}^2 \mathrm{E}[w_{j+1}^2 | \mathcal{F}_{k_j}] \xrightarrow[n \to \infty]{P} \frac{\sigma^2}{\overline{\eta}(\theta^0)} \int_0^\infty e^{2At} dt = \sigma_1^2, \tag{6.26}$$

by (6.3), Lemma 6.1 and the fact that

$$\left|\sum_{j=0}^{n} a_{k_j} \varphi_{n,\ k_j}^2 - \sum_{j=0}^{\sigma(n)} a_{k_j} \varphi_{n,\ k_j}^2\right| \le a_{k_{\sigma(n)+1}} \varphi_{n,\ k_{\sigma(n)+1}}^2 \xrightarrow[n \to \infty]{} 0, \quad a.s.$$

By condition (A4) and (6.22), we have

$$\begin{aligned} |w_{m+1}| &\le \left(\max_{\theta \in [a,\ b]} |f(\theta)| + (B_1 + B_2 L_{m+1}(b)^{\mu_0})(1 + \sum_{s=0}^{p} b_s L_{m+1}(b)^s)\right) \eta_{m+1}(b) \\ &\triangleq W_{m+1}^{(0)} \eta_{m+1}(b). \end{aligned} \tag{6.27}$$

By condition (A7) and (6.9), it is readily derived that

$$\mathrm{E}[|w_{m+1}|^3 | \mathcal{F}_{k_m}] \le \mathrm{E}[W_{m+1}^{(0)3} \eta_{m+1}(b)^3] < \infty. \tag{6.28}$$

By (6.4), it follows that

$$\sum_{j=0}^{n} a_{k_j} |\varphi_{n,\ k_j}|^3 \le \sup_n \sum_{i=0}^{n+1} a_i |\varphi_{n,\ i}|^3 < \infty.$$

For all $\varepsilon > 0$, by (6.3) and (6.28) we then have

$$\begin{aligned} &\sum_{j=0}^{n} \mathrm{E}[X_{n,\ j}^2 I_{[|X_{n,\ j}|>\varepsilon]} | \mathcal{F}_{k_j}] \\ \leq\ & \frac{1}{\varepsilon} \sum_{j=0}^{n} a_{k_j} |\varphi_{n,\ k_j}|^3 \sqrt{a_{k_j}} \mathrm{E}[|w_{j+1}|^3 | \mathcal{F}_{k_j}] \xrightarrow[n \to \infty]{} 0, \quad a.s. \end{aligned} \tag{6.29}$$

By (6.25), (6.26) and (6.29), the conditions of Theorem 9.5.2 in Chow and Teicher [9] are verified. Thus the proof of (6.24) is completed.

□

Lemma 6.3 *Suppose that conditions (A1)–(A8) are satisfied. Then*

$$\sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} \varepsilon_{k_m+i} \xrightarrow[n \to \infty]{d} \mathcal{N}(0,\ \sigma_1^2).$$

Proof. It is easy to see

$$\begin{aligned} &\sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} \varepsilon_{k_m+i} = \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f_{k_m+i} - f(\theta_{k_m+i-1})) \\ =\ & \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f_{k_m+i} - f_{k_m+i}(\theta_{k_m})) + \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f(\theta_{k_m}) - f(\theta_{k_m+i-1})) \\ &+ \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=\eta_{m+1}(\theta_{k_m})+1}^{\eta_{m+1}} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) I_{[\eta_{m+1} > \eta_{m+1}(\theta_{k_m})]} \\ &- \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=\eta_{m+1}+1}^{\eta_{m+1}(\theta_{k_m})} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) I_{[\eta_{m+1} < \eta_{m+1}(\theta_{k_m})]} \\ &+ \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}(\theta_{k_m})} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m})) \end{aligned} \tag{6.30}$$

We will prove that each of the first four terms on the right hand side of the last equality in (6.30) converges to 0 in probability. Then by Lemma 6.2 the desired result follows from (6.30).

(i). Using condition (A5), we have

$$T_{k_m+i} \leq L_{m+1}(b), \tag{6.31}$$

$$\begin{aligned} |\alpha_{k_m+j}| &\leq \sum_{l=1}^{j} \left| \frac{d\, x_{k_m+l}(\theta_{k_m+l-1})}{d\, \theta} \right| \leq \sum_{j=1}^{\eta_{m+1}(b)} \sum_{s=0}^{p} b_s (x_{k_m+j}(b))^s \\ &\leq \sum_{s=0}^{p} b_s (L_{m+1}(b))^s. \end{aligned} \tag{6.32}$$

By condition (A4), (6.31)–(6.32), it follows from (2.1)–(2.2) that

$$|\theta_{k_m+i} - \theta_{k_m}| \leq \sum_{j=1}^{i} a_{k_m+j-1} |f_{k_m+j}|$$

$$\leq \quad a_{k_m} \sum_{j=1}^{i} [(B_1 + B_2 T^{\mu_0}_{k_m+j})(|\alpha_{k_m+j}| + 1) \leq a_{k_m} W_{m+1}, \quad \forall\, 1 \leq i \leq \eta_{m+1}, \tag{6.33}$$

where

$$W_{m+1} \triangleq \eta_{m+1}(b)(B_1 + B_2 L^{\mu_0}_{m+1}(b))(1 + \sum_{s=0}^{p} b_s L^s_{m+1}(b)). \tag{6.34}$$

As in the proof (part (iv)) of Lemma 3.4 in [35], there exist random variables $\{V_{m+1},\ m \geq 0\}$ such that

$$|f_{k_m+i} - f_{k_m+i}(\theta_{k_m})| \leq a_{k_m} V_{m+1}, \quad 1 \leq i \leq \eta_{m+1}, \tag{6.35}$$

and that by condition (A7), (3.9) and the Hölder inequality the following inequality holds:

$$\sup_m \mathrm{E}[\eta_{m+1}(b) V_{m+1} | \mathcal{F}_{k_m}] < \infty. \tag{6.36}$$

Using (6.35) it follows that

$$|\sum_{m=0}^{\tau(n)} \sqrt{a_{k_m}} \varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f_{k_m+i} - f_{k_m+i}(\theta_{k_m}))| \leq \sum_{m=0}^{\tau(n)} a_{k_m} |\varphi_{n,\ k_m}| \sqrt{a_{k_m}} \eta_{m+1}(b) V_{m+1}. \tag{6.37}$$

By (6.3), (6.4), and (6.36) it is seen that

$$\mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m} |\varphi_{n,\ k_m}| \sqrt{a_{k_m}} \eta_{m+1}(b) V_{m+1}\right]$$

$$\leq \quad \mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m} |\varphi_{n,\ k_m}| \sqrt{a_{k_m}} \sup_m \{\mathrm{E}[\eta_{m+1}(b) V_{m+1} | \mathcal{F}_{k_m}]\}\right] \xrightarrow[n \to \infty]{} 0, \tag{6.38}$$

which gives

$$\sum_{m=0}^{\tau(n)} a_{k_m} |\varphi_{n,\ k_m}| \sqrt{a_{k_m}} \eta_{m+1}(b) V_{m+1} \xrightarrow[n \to \infty]{P} 0. \tag{6.39}$$

By (A6), (6.36) and Lemma 2 of [34], it is derived that

$$a_{k_m} \varphi_{n,\ k_m} \sqrt{a_{k_m}} \eta_{m+1}(b) V_{m+1} \xrightarrow[m \to \infty]{} 0, \quad a.s.,$$

which implies

$$a_{k_{\tau(n)}} \varphi_{n,\ k_{\tau(n)}} \sqrt{a_{k_{\tau(n)}}} \eta_{\tau(n)+1}(b) V_{\tau(n)+1} \xrightarrow[m \to \infty]{} 0, \quad a.s. \tag{6.40}$$

Combining (6.39) with (6.40) gives

$$\sum_{m=0}^{\sigma(n)} a_{k_m} |\varphi_{n,\ k_m}| \sqrt{a_{k_m}} \eta_{m+1}(b) V_{m+1} \xrightarrow[n \to \infty]{P} 0, \tag{6.41}$$

from which it follows that

$$
\begin{aligned}
&|\sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}}\varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f_{k_m+i} - f_{k_m+i}(\theta_{k_m}))| \\
\leq \ & \sum_{m=0}^{\sigma(n)} a_{k_m}|\varphi_{n,\ k_m}|\sqrt{a_{k_m}}\eta_{m+1}(b)V_{m+1} \xrightarrow[n\to\infty]{P} 0. \qquad (6.42)
\end{aligned}
$$

(ii). By condition (A4) and (6.33) we have

$$
|\sum_{i=1}^{\eta_{m+1}} (f(\theta_{k_m}) - f(\theta_{k_m+i-1}))| \leq B_5 \sum_{i=1}^{\eta_{m+1}} |\theta_{k_m} - \theta_{k_m+i-1}| \leq B_5 a_{k_m}\eta_{m+1}(b)W_{m+1}.
$$

Similar to (6.42), it can be shown that

$$
\begin{aligned}
&|\sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}}\varphi_{n,\ k_m} \sum_{i=1}^{\eta_{m+1}} (f(\theta_{k_m}) - f(\theta_{k_m+i-1}))| \\
\leq \ & B_5 \sum_{m=0}^{\sigma(n)} a_{k_m}|\varphi_{n,\ k_m}|\sqrt{a_{k_m}}\eta_{m+1}(b)W_{m+1} \xrightarrow[n\to\infty]{P} 0.
\end{aligned}
$$

(iii). Analogous to (6.27), by condition (A4) we have

$$
|\sum_{i=\eta_{m+1}(\theta_{k_m})+1}^{\eta_{m+1}} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m}))I_{\{\eta_{m+1}>\eta_{m+1}(\theta_{k_m})\}}| \leq \eta_{m+1}(b)W_{m+1}^{(0)} \triangleq W_{m+1}^{(3)}.
$$

By Lemma 3.3 of [35] and the Hölder inequality, it is derived that

$$
\begin{aligned}
&\mathrm{E}\left[\left|\sum_{m=0}^{\tau(n)} \sqrt{a_{k_m}}\varphi_{n,\ k_m} \sum_{i=\eta_{m+1}(\theta_{k_m})+1}^{\eta_{m+1}} (f_{k_m+i}(\theta_{k_m}) - f(\theta_{k_m}))I\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})\}\right|\right] \\
\leq \ & \mathrm{E}\left[\sum_{m=0}^{\tau(n)} \sqrt{a_{k_m}}|\varphi_{n,\ k_m}| \left(\mathrm{E}[W_{m+1}^{(3)p_2}]\right)^{1/p_2} P\{\eta_{m+1} > \eta_{m+1}(\theta_{k_m})|\mathcal{F}_{k_m}\}^{1-1/p_2}\right] \\
\leq \ & \beta_2^{1-1/p_2}\mathrm{E}\left[\sum_{m=0}^{\tau(n)} a_{k_m}|\varphi_{n,\ k_m}| \left(\mathrm{E}[W_{m+1}^{(3)p_2}]\right)^{1/p_2} a_{k_m}^{q/p_1(1-1/p_2)-1/2}\right] \xrightarrow[n\to\infty]{} 0, \qquad (6.43)
\end{aligned}
$$

as in (6.38), since $\frac{q}{p_1}(1-\frac{1}{p_2}) > \frac{1}{2}$ by condition (A7).

By the same way as in (6.42), (6.43) implies that the third term on the right hand side of the last equality in (6.30) converges to 0 in probability. Analogous to (iii), we can prove that the forth term on the right hand side of the last equality in (6.30) tends to 0 in probability.

Thus, the proof of Lemma 6.3 is completed.

□

Lemma 6.4 *If conditions (A1)-(A8) are satisfied, then*

$$\sum_{i=0}^{n}\sqrt{a_i}\varphi_{n,\ i+1}\varepsilon_{i+1}\xrightarrow[n\to\infty]{d}\mathcal{N}(0,\ \sigma_1^2).$$

Proof. By (6.2) it is seen that

$$\sum_{i=0}^{n}\sqrt{a_i}\varphi_{n,\ i+1}\varepsilon_{i+1}=\sum_{m=0}^{\sigma(n)}\sum_{i=1}^{\eta_{m+1}}\sqrt{a_{k_m+i-1}}\varphi_{n,\ k_m+i}\varepsilon_{k_m+i}$$

$$=\sum_{m=0}^{\sigma(n)}\sum_{i=1}^{\eta_{m+1}}\varphi_{n,\ k_m+i}(\sqrt{a_{k_m+i-1}}-\sqrt{a_{k_m}})\varepsilon_{k_m+i}+\sum_{m=0}^{\sigma(n)}\sqrt{a_{k_m}}\sum_{i=1}^{\eta_{m+1}}(\varphi_{n,\ k_m+i}-\varphi_{n,\ k_m})\varepsilon_{k_m+i}$$
$$+\sum_{m=0}^{\sigma(n)}\sqrt{a_{k_m}}\varphi_{n,\ k_m}\sum_{i=1}^{\eta_{m+1}}\varepsilon_{k_m+i} \tag{6.44}$$

We will prove that each of the first two terms on the right hand side of the last equality in (6.44) converges to 0 a.s. Then by Lemma 6.3, the desired result follows from (6.44).

(i). Set $\delta'=1/2$ in (6.8) we get

$$\max_{1\le i\le\eta_{m+1}}|\sqrt{a_{k_m+i-1}}-\sqrt{a_{k_m}}|\le\alpha_0^*\eta_{m+1}(b)a_{k_m}^{3/2}. \tag{6.45}$$

Similar to (6.27), we can prove

$$|\varepsilon_{k_m+i}|\le W_{m+1}^{(0)},\quad\forall\ 1\le i\le\eta_{m+1}. \tag{6.46}$$

By (6.8), (6.9) and condition (A7), similar to Lemma 3.1 in [35], we can prove

$$\lim_{m\to\infty}\sqrt{a_{k_m}}\eta_{m+1}(b)W_{m+1}^{(0)}\max_{1\le i\le\eta_{m+1}}\frac{a_{k_m}}{a_{k_m+i}}=0,\quad a.s. \tag{6.47}$$

Thus, by (6.45) it is derived that

$$\begin{aligned}&|\sum_{m=0}^{\sigma(n)}\sum_{i=1}^{\eta_{m+1}}\varphi_{n,\ k_m+i}(\sqrt{a_{k_m+i-1}}-\sqrt{a_{k_m}})\varepsilon_{k_m+i}|\\
\le\ &\alpha_0^*\sum_{m=0}^{\sigma(n)}\sum_{i=1}^{\eta_{m+1}}|\varphi_{n,\ k_m+i}|a_{k_m}^{3/2}W_{m+1}^{(0)}\eta_{m+1}(b)\\
\le\ &\alpha_0^*\sum_{m=0}^{\sigma(n)}\sum_{i=1}^{\eta_{m+1}}a_{k_m+i}|\varphi_{n,\ k_m+i}|\sqrt{a_{k_m}}W_{m+1}^{(0)}\eta_{m+1}(b)\max_{1\le i\le\eta_{m+1}}\frac{a_{k_m}}{a_{k_m+i}}\\
&\xrightarrow[n\to\infty]{}0,\quad a.s.,\end{aligned} \tag{6.48}$$

via (6.3), (6.4) and (6.47).

(ii). Since $\lim_n A_n=A$, there are constants α_2^* and m_0^* such that

$$|A_n|\le\alpha_2^*,\quad\forall\ n\ge1;$$
$$0<\varphi_{n,\ k_m+j-1}\le\varphi_{n,\ k_m+i},\quad\forall\ 1\le j\le i,\ 1\le i\le\eta_{m+1},\ m\ge m_0^*.$$

Then by (6.2) and (6.46) we have

$$
\begin{aligned}
& \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \sum_{i=1}^{\eta_{m+1}} (\varphi_{n,\ k_m+i} - \varphi_{n,\ k_m}) \varepsilon_{k_m+i} \\
\leq\ & \sum_{m=0}^{\sigma(n)} \sqrt{a_{k_m}} \sum_{i=1}^{\eta_{m+1}} \sum_{j=1}^{i} a_{k_m+j-1} |A_{k_m+j-1} \varphi_{n,\ k_m+j}| W_{m+1}^{(0)} \\
\leq\ & \alpha_2^* \sum_{m=0}^{m_0^*-1} a_{k_m}^{3/2} \sum_{i=1}^{\eta_{m+1}} \sum_{j=1}^{i} |\varphi_{n,\ k_m+j}| W_{m+1}^{(0)} \\
& + \alpha_2^* \sum_{m=m_0^*}^{\sigma(n)} \sum_{i=1}^{\eta_{m+1}} a_{k_m+i-1} |\varphi_{n,\ k_m+i}| \sqrt{a_{k_m}} \eta_{m+1}(b) W_{m+1}^{(0)} \max_{1 \leq i \leq \eta_{m+1}} \frac{a_{k_m}}{a_{k_m+i-1}} \\
& \xrightarrow[n \to \infty]{} 0, \quad a.s.,
\end{aligned}
$$

via a similar argument as in (6.48).

□

Proof of Theorem 2.1:

Under conditions (A1)–(A7), from Theorem 2.2 of [35] we have $|\theta_n - \theta^0| = o(a_n^{1/4})$ a.s., as $n \to \infty$. After a finite number of steps, algorithm (2.1) will become the usual Robbins-Monro algorithm. That is, there is a constant n_0 (may depend on sample path) such that

$$\theta_{n+1} = \theta_n - a_n(f(\theta_n) + \varepsilon_{n+1}), \quad \forall\ n \geq n_0.$$

The standard treatment leads to

$$
\begin{aligned}
\frac{\theta_{n+1} - \theta^0}{\sqrt{a_{n+1}}} &= (1 + a_n A_n) \frac{\theta_n - \theta^0}{\sqrt{a_n}} - \sqrt{a_n}(1 + \alpha a_n/2 + o(a_n))(\varepsilon_{n+1} + O(|\theta_n - \theta^0|^2)) \\
&= \varphi_{n,\ n_0} \frac{\theta_{n_0} - \theta^0}{\sqrt{a_{n_0}}} - \sum_{i=n_0}^{n} \varphi_{n,\ i+1} \sqrt{a_i} \varepsilon_{i+1} - \sum_{i=n_0}^{n} \varphi_{n,\ i+1} a_i \sqrt{a_i} (\alpha/2 + o(1)) \varepsilon_{i+1} \\
&\quad - \sum_{i=n_0}^{n} \varphi_{n,\ i+1} \sqrt{a_i} (1 + \alpha a_i/2 + o(a_i)) O(|\theta_i - \theta^0|^2),
\end{aligned} \tag{6.49}
$$

where $A_n = -M_1 + \alpha/2 + o(1) \xrightarrow[n \to \infty]{} A \triangleq -M_1 + \alpha/2 < 0$ and $\varphi_{n,\ i},\ n \geq 0,\ i \geq 0$ are defined by (6.2). By (6.46) and (6.47), it is seen that

$$\max_{1 \leq i \leq \eta_{m+1}} \sqrt{a_{k_m+i-1}} |\varepsilon_{k_m+i}| \leq \sqrt{a_{k_m}} W_{m+1}^{(0)} \xrightarrow[m \to \infty]{} 0, \quad a.s.,$$

which yields that

$$\lim_{n \to \infty} \sqrt{a_n} \varepsilon_{n+1} = 0, \quad a.s. \tag{6.50}$$

By (6.3), (6.4) and (6.50), it follows that

$$\sum_{i=n_0}^{n} \varphi_{n,\ i+1} a_i \sqrt{a_i} (\alpha/2 + o(1)) \varepsilon_{i+1} \xrightarrow[n \to \infty]{} 0, \quad a.s. \tag{6.51}$$

Similarly, we can prove that the first and the last terms on the right hand side hand of the last equality in (6.49) converge to 0 a.s. Then by Lemma 6.4, the theorem follows from (6.49).

□

Proof of Theorem 3.1:

By (6.49), to prove the theorem the key step is to prove

$$\sum_{i=0}^{n} \sqrt{a_i} \varphi_{n,\ i+1} \varepsilon_{i+1} \xrightarrow[n \to \infty]{d} \mathcal{N}(0,\ \sigma_2^2), \tag{6.52}$$

where $\varepsilon_{n+1} = f_{n+1} - f(\theta_n)$ is the observation noise with f_{n+1} defined by (3.1), $\forall\ n \geq 0$.

In order to make the result of Theorem 2.1 applicable to the present setting, as in [34] and [35], we introduce the following transformations

$$\mu(n) = [\frac{n}{L}], \quad \widetilde{\theta}_n = \theta_{\mu(n)}, \quad \widetilde{a}_n = a_{\mu(n)}; \tag{6.53}$$

$$\beta_{(n-1)L+j} = \beta_{n,\ j}, \quad j = 0,\ 1,\ ...,\ L-1; \tag{6.54}$$

$$\widetilde{\varepsilon}_n = J_t(T_n,\ \widetilde{\theta}_{n-1})\beta_n + J_\theta(T_n,\ \widetilde{\theta}_{n-1}) - f(\widetilde{\theta}_{n-1}); \tag{6.55}$$

$$\widetilde{\varphi}_{n,\ i} = \varphi_{\mu(n-1),\ \mu(i-1)+1},\ \forall\ i \geq 1;\ \widetilde{\varphi}_{n,\ j} = \widetilde{\varphi}_{n,\ 0},\ \forall\ j < 0; \quad \forall\ n \geq 1. \tag{6.56}$$

Notice that $\widetilde{\theta}_n$ is the actual service parameter of the $(n+1)$-th customer. From (3.2) and (6.54) we have

$$\beta_{n+1} = \beta_n I_{[Q_n \geq 1]} + \frac{d\, x_{n+1}(\widetilde{\theta}_n)}{d\, \theta}, \quad \forall\ n \geq 0. \tag{6.57}$$

By (6.2), it follows from (6.56) that

$$\widetilde{\varphi}_{n,\ j} = 0, \quad \forall\ j \geq n + L. \tag{6.58}$$

By (3.1), (6.53)–(6.56) it is seen that

$$\sum_{i=0}^{n} \sqrt{a_i} \varphi_{n,\ i+1} \varepsilon_{i+1} = \sum_{i=0}^{n} \sqrt{a_i} \varphi_{n,\ i+1} (f_{i+1} - f(\theta_i))$$

$$= \frac{1}{L} \sum_{i=0}^{n} \sqrt{a_i} \varphi_{n,\ i+1} \sum_{l=1}^{L} \widetilde{\varepsilon}_{iL+l} = \frac{1}{L} \sum_{i=1}^{(n+1)L} \sqrt{\widetilde{a}_{i-1}} \widetilde{\varphi}_{(n+1)L,\ i} \widetilde{\varepsilon}_i, \tag{6.59}$$

which means that in order to prove (6.52), it is equivalent to prove that

$$\frac{1}{L} \sum_{i=1}^{(n+1)L} \sqrt{\widetilde{a}_{i-1}} \widetilde{\varphi}_{(n+1)L,\ i} \widetilde{\varepsilon}_i \xrightarrow[n \to \infty]{d} \mathcal{N}(0,\ \sigma_2^2). \tag{6.60}$$

The proof of (6.60) follows essentially the same avenue as used in Lemmas 6.1–6.4 if we replace θ_n, ε_n, a_n, α_n, $\varphi_{n,\ i}$ by $\widetilde{\theta}_n$, $\widetilde{\varepsilon}_n$, $\widetilde{a}_n$, β_n, $\widetilde{\varphi}_{(n+1)L,\ i}$ respectively. We give the outline of the proof.

By (6.5), (6.53) and (6.56) it is easy to see that

$$\sum_{i=1}^{(n+1)L} \widetilde{a}_{i-1}\widetilde{\varphi}^2_{(n+1)L,\ i} = \sum_{i=1}^{(n+1)L} a_{\mu(i-1)}\varphi^2_{\mu((n+1)L-1),\ \mu(i-1)+1}$$

$$= \quad L\sum_{i=0}^{n} a_i\varphi^2_{n,\ i+1} \xrightarrow[n\to\infty]{} L\int_0^\infty e^{2At}dt, \tag{6.61}$$

which yields

$$\sum_{m=0}^{\sigma((n+1)L)}\sum_{i=1}^{\eta_{m+1}} \widetilde{a}_{k_m+i-1}\widetilde{\varphi}^2_{(n+1)L,\ k_m+i} \xrightarrow[n\to\infty]{} L\int_0^\infty e^{2At}dt, \quad a.s. \tag{6.62}$$

Before applying Lemmas 6.1–6.4 to the present setting, let us verify that similar results for $\widetilde{a}_n$ and $\widetilde{\theta}_n$ hold as in (6.8) and (6.33).

We first check (6.33) for $\widetilde{\theta}_n$. If $\mu(k_m+i) = \mu(k_m)$, then $\widetilde{\theta}_{k_m+i} = \widetilde{\theta}_{k_m}$. Otherwise, let $r_1 \triangleq \mu(k_m) < \mu(k_m+i) \triangleq r_2$. Then by condition (A4), (6.31)–(6.32), it follows from (3.1) and (6.57) that

$$|\widetilde{\theta}_{k_m+i} - \widetilde{\theta}_{k_m}| = |\theta_{r_2} - \theta_{r_1}| \le \sum_{j=r_1}^{r_2} a_j|f_{j+1}|$$

$$\le \quad \widetilde{a}_{k_m}\sum_{j=r_1}^{r_2-1}\frac{1}{L}\sum_{l=1}^{L}|J_t(T_{jL+l},\ \theta_j)\beta_{jL+l} + J_\theta(T_{jL+l},\ \theta_j)|$$

$$\le \quad \widetilde{a}_{k_m}\frac{1}{L}\sum_{j=r_1L+1}^{Lr_2}|J_t(T_j,\ \widetilde{\theta}_{j-1})\beta_j + J_\theta(T_j,\ \widetilde{\theta}_{j-1})|$$

$$\le \quad \widetilde{a}_{k_m}\frac{1}{L}\sum_{j=k_m-L+1}^{k_m+\eta_{m+1}}(B_1 + B_2T_j^{\mu_0})(1+|\beta_j|)$$

$$\le \quad \widetilde{a}_{k_{m-L}}\widetilde{W}_{m+1}, \qquad \forall\ 1\le i\le \eta_{m+1}, \tag{6.63}$$

where

$$\widetilde{W}_{m+1} = \frac{1}{L}\sum_{j=0}^{L}\eta_{m-j+1}(b)(B_1 + B_2L_{m-j+1}(b)^{\mu_0})(1+\sum_{s=0}^{p}b_sL_{m-j+1}(b)^s). \tag{6.64}$$

Repeating the above arguments, we also have

$$|\widetilde{a}^{1-\delta'}_{k_m+i} - \widetilde{a}^{1-\delta'}_{k_{m-L}}| \le \alpha_0^*(1+\sum_{j=0}^{L}\eta_{m-j+1}(b))\widetilde{a}^{2-\delta'}_{k_{m-L}}, \quad \forall\ \delta'\in[0,1),\ 1\le i\le\eta_{m+1}. \tag{6.65}$$

Along the same lines of Lemma 6.1, by (6.62) it can be shown that

$$\sum_{m=0}^{\sigma(nL+L)} \widetilde{a}_{k_{m-L}}\widetilde{\varphi}^2_{nL+L,\ k_{m-L}} \xrightarrow[n\to\infty]{} \frac{L}{\overline{\eta}(\theta^0)}\int_0^\infty e^{2At}dt. \tag{6.66}$$

By (6.66) and Theorem 9.5.2 in Chow and Teicher [9], analogous to Lemma 6.2, it is derived that

$$\frac{1}{L}\sum_{m=0}^{\sigma(nL+L)}\sqrt{\widetilde{a}_{k_{m-L}}}\widetilde{\varphi}_{nL+L,\ k_{m-L}}\sum_{i=1}^{\eta_{m+1}(\widetilde{\theta}_{k_m})}(f_{k_m+i}(\widetilde{\theta}_{k_m})-f(\widetilde{\theta}_{k_m}))\xrightarrow[n\to\infty]{d}\mathcal{N}(0,\ \sigma_2^2). \tag{6.67}$$

Similar to Lemma 6.3, by (6.67) we can prove

$$\frac{1}{L}\sum_{m=0}^{\sigma(nL+L)}\sqrt{\widetilde{a}_{k_m-L}}\widetilde{\varphi}_{nL+L,\ k_{m-L}}\sum_{i=1}^{\eta_{m+1}}\widetilde{\varepsilon}_{k_m+i}\xrightarrow[n\to\infty]{d}\mathcal{N}(0,\ \sigma_2^2), \tag{6.68}$$

which yields (6.60) by the same way as in Lemma 6.4. This completes the proof of the theorem.

□

References

[1] S. Asmussen, *Applied Probability and Queues*, New York: John Wiley and Sons, 1987.

[2] H. F. Chen,"Asymptotic efficient stochastic approximation,"*Stochastics and Stochastics Reports*, vol.45, pp.1–16, 1993.

[3] H. F. Chen, "Stochastic approximation and its new applications," in *Proc. 1994 Hong Kong International Workshop on New Directions in Control and Manufacturing*, pp.2–12, 1994.

[4] H. F. Chen and Y. M. Zhu, *Stochastic Approximation*, Shanghai Scientific and Technological Publishers, 1996.

[5] E. K. P. Chong and P. J. Ramadge, "Convergence of recursive optimization algorithms using infinitesimal perturbation analysis estimates," *Discrete Event Dynamic Systems: Theory and Appl.*, vol.1, pp. 339–372, 1992.

[6] E. K. P. Chong and P. J. Ramadge, "Optimization of queues using an infinitesimal perturbation analysis-based stochastic algorithm with general updates times," *SIAM J. Contr. Optim.*, vol.31, pp. 698–732, 1993.

[7] E. K. P. Chong and P. J. Ramadge, "Stochastic optimization of regenerative systems using infinitesimal perturbation analysis," *IEEE Trans. Automat. Contr.*, vol. AC-39, no.7, pp. 1400-1410, 1994.

[8] Y. S. Chow, "Local convergence of martingale and the law of large numbers," *Ann. Math. Stat.*, vol.36, pp. 552–558, 1965.

[9] Y. S. Chow and H. Teicher, *Probability Theory : Independence, Interchangeability, Martingale*, 1988, Second Edition, Springer-Verlag.

[10] M. C. Fu, "Convergence of a stochastic approximation algorithm for the GI/G/1 queue using infinitesimal perturbation analysis," *J. of Optim. Theory and Appl.*, vol.65, pp. 149–160, 1990.

[11] M. C. Fu and Y. C. Ho, "Using perturbation analysis for gradient estimation, averaging, and updating in a stochastic approximation algorithm," *Proc. of the Winter Simulation Conference*, pp. 509-517, 1988.

[12] P. Glasserman, *Gradient Estimation via Perturbation Analysis*, Kluwer Academic Pub., Boston, 1991.

[13] A. Gut, *Stopped Random Walks: Limit Theorems and Applications*, Springer-Verlag, 1988.

[14] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application.* New York: Acad. Press, 1980.

[15] Y. C. Ho and X. R. Cao, "Perturbation analysis and optimization of queueing networks," *J. Optimiz. Theory Appl.*, vol.40, pp.559-582, 1983.

[16] Y. C. Ho and X. R. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*, Kluwer Academic Pub., Boston, 1991.

[17] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. New York: Springer, 1978.

[18] H. J. Kushner and F. J. Vázquez-Abad, "Stochastic approximation methods for systems of interest over an infinite horizon," *SIAM J. on Control and Optim.*, vol.34, pp.712-756, 1996.

[19] P. L'Ecuyer, "A Unified View of the IPA, SF, and LR Gradient Estimation Techniques," *Management Science*, Vol. 36, pp. 1364–1383, 1990.

[20] P. L'Ecuyer and P. W. Glynn, "Stochastic optimization by simulation: convergence proofs for the GI/G/1 queue in steady-state,"*Management Sci.*, vol.40, pp.1562-1578, 1994.

[21] P. L'Ecuyer, N. Giroux, and P. W. Glynn, "Stochastic optimization by simulation: numerical experiments with the M/M/1 queue in steady-state," *Management Sci.*, vol.40, pp.1245–1261, 1994.

[22] P. L'Ecuyer and G. Yin, "Budget-dependent convergence rate of stochastic approximation," *SIAM Journal on Optimization*, vol 8, pp. 217–247, 1998.

[23] M. S. Meketon, "Optimization in simulation: a survey of recent results," *Proc. of the 1987 Winter Simulation Conference*, pp. 58–67, 1987.

[24] M. B. Nevel'son and R. S. Hasminsky, *Stochastic Approximation and Recursive Estimation*, Translated in *Amer. Math. Soc.*, Transl. Monographs, vol.24, Providence, 1972.

[25] E. Nummelin, "Regeneration in tandem queues". *Adv. Appl. Prob.*, vol.13, pp.221–230, 1981.

[26] B. T. Polyak, "New method of stochastic approximation type procedure,"*Automation and Remote Control*, vol.51, pp.937–946, 1990.

[27] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol.22, pp. 400-407, 1951.

[28] K. Sigman, "Notes on the stability of closed queueing networks," *J. Appl. Probab.*, vol.26, pp.678-682, 1989; "Corrections," J. Appl. Probab., vol. 27, p.735, 1990.

[29] K. Sigman, "The stability of open queueing networks," *Stoch. Processes and their Appls.*, vol.35, pp.11-25, 1990.

[30] W. F. Stout, *Almost Sure Convergence*, New York: Academic Press, 1974.

[31] R. Suri, "Perturbation analysis: the state of the art and research issues explained via the GI/G/1 queue," *Proc. IEEE*, vol. 77, pp. 114–137, 1989.

[32] R. Suri and Y. T. Leung, "Single run optimization of discrete event simulations — An empirical study using the M/M/1 queue," *IIE Transactions*, vol.21, pp. 35-49, 1989.

[33] R. Suri and M. A. Zazanis, "Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue," *Management Sci.*, vol.34, pp.39-64, 1988.

[34] Q. Y. Tang and H. F. Chen, "Convergence of perturbation analysis based optimization algorithm with fixed number of customers period," *Discrete Event Dynamic Systems: Theory and Appl.*, vol.4, pp.359-375, 1994.

[35] Q. Y. Tang, H. F. Chen, and Z. J. Han, "Convergence rates of Perturbation-Analysis-Robbins-Monro-Single-Run algorithms for single server queues," *IEEE Trans. on Autom. Contr.*, vol. 42, no. 10, pp. 1442-1447, 1997.

[36] H. Thorisson, "The queue $GI/G/1$: finite moments of the cycle variables and uniform rates of convergence," *Stoch. Processes and their Appl.*, vol.19, pp. 85-99, 1985.

[37] R. Wolff, *Stochastic Modeling and the Theory of Queues*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[38] G. Yin, "On extension of Polyak's averaging approach to stochastic approximation," *Stochastics and Stochastics Reports*, vol. 36, pp.245–264, 1991.