On the Convergence of a Two Timescale Stochastic Approximation Algorithm for Optimal Observer Trajectory Planning

Sumeetpal S. Singh[a], Nikolaos Kantas[a]

Ba-Ngu Vo[b], Arnaud Doucet[a], Robin J. Evans[b]

[a]Signal Processing Group, Department of Engineering
University of Cambridge
Trumpington Street, Cambridge CB2 1PZ, UK
{sss40,nk234,ad2}@eng.cam.ac.uk

[b]Department of Electrical and Electronic Engineering
University of Melbourne
VIC 3010, Australia
{b.vo,r.evans}@ee.mu.oz.au

March 29, 2005

Abstract

Sensor scheduling has been a topic of interest to the target tracking community for some years now and more recently, it has enjoyed fresh impetus with the current importance and popularity of applications in Sensor Networks and Robotics. The sensor scheduling problem can be formulated as a controlled Hidden Markov Model. In the paper [7], we addressed precisely this problem and considered the case in which the state, observation and action spaces are continuous. We presented a novel simulation-based method that used a two timescale stochastic approximation algorithm to find optimal actions. In this report, we study the convergence of the proposed stochastic approximation algorithm under general assumptions and for the specific observer trajectory planning application of [7].

1 The Main Algorithm and its Convergence

The simulation-based method proposed in [7] to solve the sensor scheduling problem is a two timescale stochastic approximation (SA) algorithm. We commence the report by presenting this SA algorithm in its general form and study its convergence. Section 2 presents a small variant of the SA algorithm studied here which is more suitable from the observer trajectory planning (OTP) application. (Note that OTP may be viewed a sensor scheduling problem.) We show that the convergence results carry over easily. Finally, the OTP problem of [7] is presented in Section 3 and the convergence of two timescale SA for this application is established.

Notation 1 *The notation that is used in the report is now outlined. The norm of a scalar, vector or matrix is denoted by* $|\cdot|$. *For a vector* b, $|b|$ *denotes the vector 2-norm* $\sqrt{\sum_i |b(i)|^2}$. *For a matrix* A, $|A|$ *denotes the matrix 2-norm,* $\max_{b:|b|\neq 0} \frac{|Ab|}{|b|}$. *For convenience, we also denote a vector* $b \in R^n$ *by* $b = [b(i)]_{i=1,\ldots,n}$, *or the i-th component of a vector by* $[b]_i$. *For scalars* $a_{j,i}$, $j = 1,\ldots,m$, $i = 1,\ldots,n$, *let* $\left[[a_{j,i}]_{j=1,\ldots,m}\right]_{i=1,\ldots,n}$ *denote the stacked vector* $[a_{1,1},\ldots,a_{m,1},\ldots,a_{1,n},\ldots,a_{m,n}]^T$*.) For a vector* b, *let diag(b) denote the diagonal matrix formed from* b. *For a function* $f : R^n \to R$, *we denote* $(\partial f/\partial z(i))(z)$ *by* $\nabla_{z(i)} f(z)$ *and* $\nabla f(z) = [\nabla_{z(1)} f(z), \cdots, \nabla_{z(n)} f(z)]^T$. *For the vector valued function* $F = [F_1, \cdots, F_n]^T : R^n \to R^n$, *let* ∇F *denote the matrix* $[\nabla F_1, \cdots, \nabla F_n]$. *For real-valued integrable functions* f *and* g, *let* $\langle f, g\rangle$ *denote* $\int f(x)g(x)dx$. *For a random variable* X *and probability* P, $X \sim P$ *implies the law of* X *is* P. $\mathbf{1}$ *denotes the vector with elements 1.*

Consider the following two timescale SA algorithm, which is the main SA algorithm of this report,

$$\theta_{k+1} = \theta_k - \alpha_{k+1}\Gamma(b_k)\left(h_{1,\theta_k}(\omega_{k+1}) + h_{2,\theta_k}(\omega_{k+1}) - S_{\theta_k}(\omega_{k+1})b_k\right), \quad (1)$$

$$b_{k+1} = b_k - \beta_{k+1}\left(S^2_{\theta_k}(\omega_{k+1})b_k - S_{\theta_k}(\omega_{k+1})h_{1,\theta_k}(\omega_{k+1})\right), \quad (2)$$

$$\omega_{k+1} \sim P_{\theta_k}, \qquad k \geq 0, \quad (3)$$

where $b_k, \theta_k \in R^d$. For each $\theta \in R^d$, functions $h_{1,\theta}$ and $h_{2,\theta}$ are R^d-valued, i.e., $h_{1,\theta}, h_{2,\theta} : \Omega \to R^d$. Likewise, for each $\theta \in R^d$, S_θ is a $d \times d$ diagonal matrix valued function, i.e., $S_\theta : \Omega \to R^{d\times d}$. The step-sizes α_k and β_k are non-negative scalars. The scalar valued function $\Gamma : R^d \to (0,\infty)$ is given as follows,

$$\Gamma(b) = \frac{C}{1+|b|}, \quad (4)$$

where $C > 0$ is a constant. The function Γ is needed to ensure that the iterates b_k remain bounded almost surely (see Assumption 3 below). Define the σ-algebra

$$\mathcal{F}_k = \sigma\{\theta_0\}, \qquad \mathcal{F}_k = \sigma\{\theta_0, \omega_1, \ldots, \omega_k\}, \qquad k \geq 1,$$

and let $E_k(\cdot)$ denote $E(\cdot|\mathcal{F}_k)$. It follows that θ_k and b_k are $\mathcal{F}_k$-measurable. For each $\theta \in R^d$, we assume the function S_θ satisfies

$$E_{\omega\sim P_\theta}(S_\theta(\omega)) = 0.$$

For each $\theta \in R^d$, define

$$\bar{h}_i(\theta) = E_{\omega\sim P_\theta}(h_{i,\theta}(\omega)), \qquad i = 1, 2, \quad (5)$$

$$h_\theta(\omega) = h_{1,\theta}(\omega) + h_{2,\theta}(\omega), \qquad \omega \in \Omega, \quad (6)$$

$$\bar{h}(\theta) = \bar{h}_1(\theta) + \bar{h}_2(\theta). \quad (7)$$

Additionally, we assume that $\bar{h}$ itself is a gradient of some performance criterion J, i.e.,

$$\bar{h}(\theta) = \nabla J(\theta).$$

Thus when ω is sampled according to P_θ, $h_\theta(\omega)$ is an unbiased estimate of $\nabla J(\theta)$.

The convergence of a two timescale SA algorithm similar to (1)-(3) was studied in [5]. We may write the slow time-scale process in a more general form as

$$\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}. \tag{8}$$

If the parameter θ_k did not change, say $\theta_k = \theta$ for all k, the process $\{b_k\}$ would converge to some $\bar{b}(\theta)$. When θ_k varies slowly, we would like the process $\{b_k\}$ to track $\bar{b}(\theta_k)$. Under certain regularity assumptions on the process $\{H_k\}$ (see Section 1.1 below), it can be shown that the process $\{b_k\}$ converges in the sense

$$\lim_k \left| b_k - E_{\omega \sim P_{\theta_k}} \left(S^2_{\theta_k}(\omega)\right)^{-1} E_{\omega \sim P_{\theta_k}} \left(S_{\theta_k}(\omega) h_{1,\theta_k}(\omega)\right) \right| = 0.$$

As for the convergence of $\{\theta_k\}$, we will use the line of proof in [1], which are based on the classical arguments of [6] with slightly modified assumptions, to show $\liminf_k \left|\bar{h}(\theta_k)\right| = 0$.

1.1 Convergence of the fast timescale

The assumptions below to establish the convergence of process $\{b_k\}$ are essentially the same as in [5] but with some omissions. These are due to the Markov structure of ω_{k+1} in [5], i.e., ω_{k+1} depends on θ_k and ω_k, while in our case, there is no dependence on ω_k.

Assumption 1 The step-size sequences $\{\alpha_k\}$ and $\{\beta_k\}$ satisfy

$$\begin{aligned} \alpha_k, \beta_k &\geq 0, \qquad \sum_k \alpha_k = \sum_k \beta_k = \infty, \\ \sum_k \alpha_k^2 &< \infty, \qquad \sum_k \beta_k^2 < \infty, \\ \sum_k \left(\frac{\alpha_k}{\beta_k}\right)^p &< \infty, \end{aligned}$$

for some $p > 0$.

Assumption 2 Define the following functions,

$$\begin{aligned} \overline{S^2}(\theta) &= E_{\omega \sim P_\theta} \left(S^2_\theta(\omega)\right), \\ \overline{S \times h_1}(\theta) &= E_{\omega \sim P_\theta} \left(S_\theta(\omega) h_{1,\theta}(\omega)\right). \end{aligned}$$

(a) There exists some constant C such that for all $\theta \in R^d$, we have

$$\max\left(\left|\overline{S^2}(\theta)\right|, \left|\overline{S \times h_1}(\theta)\right|\right) \leq C.$$

(b) There exists some constant C such that for all $\theta, \theta' \in R^d$, we have

$$\max\left(\left|\overline{S^2}(\theta) - \overline{S^2}(\theta')\right|, \left|\overline{S \times h_1}(\theta) - \overline{S \times h_1}(\theta')\right|\right) \leq C \left|\theta - \theta'\right|.$$

(c) For each $p > 0$, there exists a constant $C_p > 0$ such that almost surely[1],

$$\begin{aligned} \sup_k E_{\omega_{k+1} \sim P_{\theta_k}} \left(\left|S^2_{\theta_k}(\omega_{k+1})\right|^p\right) &< C_p, \\ \sup_k E_{\omega_{k+1} \sim P_{\theta_k}} \left(\left|S_{\theta_k}(\omega_{k+1}) h_{1,\theta_k}(\omega_{k+1})\right|^p\right) &< C_p. \end{aligned}$$

[1]The proof actually requires the following weaker set of conditions [5]:

$$\begin{aligned} \sup_k E\left(\left|S^2_{\theta_k}(\omega_{k+1})\right|^p\right) &< C_p, \\ \sup_k E\left(\left|S_{\theta_k}(\omega_{k+1}) h_{1,\theta_k}(\omega_{k+1})\right|^p\right) &< C_p. \end{aligned}$$

However, in Section 4.1.1, we verify the stated stronger assumptions for the OTP problem.

Assumption 3 Re-writing the iteration for θ_k as $\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}$, we require

$$\sup_k E\left(|H_k|^p\right) < \infty$$

for all $p > 0$.

Assumption 4 (Uniform positive definiteness) There exists some constant $a > 0$ such that for all $b, \theta \in R^d$,

$$b^T \overline{S^2}(\theta) b \geq a|b|^2.$$

The proof of the following result is available in [5].

Lemma 2 *[5, Theorem 7] If Assumptions 1-4 are satisfied then, almost surely,* $\sup_k |b_k| < \infty$ *and* $\lim_k \left| b_k - \overline{S^2}(\theta_k)^{-1} \overline{S \times h_1}(\theta_k) \right| = 0$.

Remark 3 *The statement of the above lemma concerning the convergence of* $\left| b_k - \overline{S^2}(\theta_k)^{-1} \overline{S \times h_1}(\theta_k) \right|$ *is general in the sense that the only restriction imposed on the recursion for* θ_k *is that it should be able to be written in the form* $\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}$ *with* H_k *satisfying Assumption 3.*

Remark 4 *Concerning Assumption 3, strictly speaking, we only need* $\sup_k E\left(|H_k|^p\right) < \infty$ *to be satisfied for the value of* p *for which* $\sum_k \left(\frac{\alpha_k}{\beta_k}\right)^p < \infty$. *However, from a practical point of view, it is not much more difficult to verify Assumption 3 for all* $p > 0$. *Also, the choice of step-sizes are typically*

$$\alpha_k = k^{-\alpha}, \qquad \beta_k = k^{-\beta},$$

where constants α *and* β *are allowed to assume values from the range* $\alpha > \beta > 0.5$. *Thus,* $\sum_k \left(\frac{\alpha_k}{\beta_k}\right)^p < \infty$ *may only be satisfied for a large positive* p.

1.2 Convergence of the slow timescale

The proof of the convergence of $\nabla J(\theta_k)$ below is based on the approach in [1]. Before we present the analysis, we require the following two lemmas from [1].

Lemma 5 *[1, Lemma 1] Let* Y_k, W_k *and* Z_k *be three real valued sequences such that* W_k *is non-negative for all* k. *Assume that*

$$Y_{k+1} \leq Y_k - W_k + Z_k,$$

$k \geq 0$, *and* $\lim_{p \to \infty} \sum_{k=0}^{p} Z_k$ *exists. Then, either* $Y_k \to -\infty$ *or else* Y_k *converges to a finite value and* $\sum_{k=0}^{\infty} W_k < \infty$.

Lemma 6 *[1] Let* $f : R^d \to R$ *be a continuously differentiable function satisfying for some constant* C,

$$\left|\nabla f(\theta) - \nabla f(\theta')\right| \leq C \left|\theta - \theta'\right|,$$

for all $\theta', \theta \in R^d$. *Then,*

$$f(\theta') - f(\theta) \leq (\theta' - \theta)^T \nabla f(\theta) + \frac{C}{2} \left|\theta' - \theta\right|^2.$$

We now present the line of arguments to establish the convergence of $\nabla J(\theta_k)$ and then formally state the result as a lemma.

In (8) and Assumption 3 above, we wrote $\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}$, where

$$\begin{aligned} H_{k+1} &= -\Gamma(b_k)\left(h_{1,\theta_k}(\omega_{k+1}) + h_{2,\theta_k}(\omega_{k+1}) - S_{\theta_k}(\omega_{k+1}) b_k\right) \\ &= -\Gamma(b_k)\bar{h}(\theta_k) + W_{k+1}, \end{aligned}$$

and

$$W_{k+1} = -\Gamma(b_k)\left(h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right) -\Gamma(b_k)\left(h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right) + \Gamma(b_k)S_{\theta_k}(\omega_{k+1})b_k.$$

Note that $\{W_k\}$ satisfies $E_k\left(W_{k+1}\right) = 0$. We assume $\bar{h}(\theta)$ is bounded. Also, assume $\bar{h}(\theta)$ $(= \nabla J(\theta))$ satisfies for some constant C and all θ, θ',

$$|\bar{h}(\theta) - \bar{h}(\theta')| < C\,|\theta - \theta'|\,.$$

By Lemma 6,

$$\begin{aligned} J(\theta_{k+1}) &\le J(\theta_k) + (\theta_{k+1} - \theta_k)^T\,\bar{h}(\theta_k) + \frac{C}{2}\,|\theta_{k+1} - \theta_k|^2 \\ &= J(\theta_k) - \alpha_{k+1}\Gamma(b_k)\left|\bar{h}(\theta_k)\right|^2 + \alpha_{k+1}W_{k+1}^T\bar{h}(\theta_k) + \frac{C}{2}\alpha_{k+1}^2\,|H_{k+1}|^2\,. \end{aligned}$$

Assumption 3 asserts $\sup_k E\left(|H_k|^2\right) < \infty$. Thus,

$$E\left\{\lim_{p\to\infty}\sum_{k=1}^{p}\alpha_k^2\,|H_k|^2\right\} = \lim_{p\to\infty}\sum_{k=1}^{p}\alpha_k^2 E\left\{|H_k|^2\right\} < \infty,$$

which implies

$$\sum_{k=1}^{\infty}\alpha_k^2\,|H_k|^2$$

exists almost surely. If we can also assert

$$\sum_{k=0}^{\infty}\alpha_{k+1}W_{k+1}^T\bar{h}(\theta_k)$$

exists almost surely then, we can invoke Lemma 5 below to conclude that either $J(\theta_k) \to -\infty$ or else $J(\theta_k)$ converges to a finite value and $\sum_{k=0}^{\infty}\alpha_{k+1}\Gamma(b_k)\left|\bar{h}(\theta_k)\right|^2 < \infty$, where the remark holds almost surely. Since J is bounded below, $\sum_{k=0}^{\infty}\alpha_{k+1}\Gamma(b_k)\left|\bar{h}(\theta_k)\right|^2 < \infty$ almost surely. Also, since $\sup_k |b_k| < \infty$ almost surely, $\Gamma(b_k)$ is bounded below and $\liminf_k \left|\bar{h}(\theta_k)\right| = 0$. Note that if $\liminf_k \left|\bar{h}(\theta_k)\right| > 0$ then, for large enough T and some $\epsilon > 0$, $\inf_{k\ge T}\left|\bar{h}(\theta_k)\right| > \epsilon$, which would contradict the fact that $\sum_{k=0}^{\infty}\alpha_{k+1}\Gamma(b_k)\left|\bar{h}(\theta_k)\right|^2 < \infty$.

To assert $\sum_{k=0}^{\infty}\alpha_{k+1}W_{k+1}^T\bar{h}(\theta_k)$ exists almost surely, we will need the MCT (Martingale Convergence Theorem). Note that

$$\begin{aligned} |W_{k+1}|^2 &= \Big|-\Gamma(b_k)\left(h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right) \\ &\quad -\Gamma(b_k)\left(h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right) + \Gamma(b_k)S_{\theta_k}(\omega_{k+1})b_k\Big|^2 \\ &\le \Big(\left|h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right| \\ &\quad + \left|h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right| + |\Gamma(b_k)S_{\theta_k}(\omega_{k+1})b_k|\Big)^2 \\ &\le 2\left|h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right|^2 \\ &\quad +4\left|h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right|^2 + 4\,|\Gamma(b_k)S_{\theta_k}(\omega_{k+1})b_k|^2\,, \end{aligned}$$

where in the last line we have used the inequality (see A.N. Shiryaev, Probability, 2nd Ed., pg.194): if $a, b > 0$ and $p \ge 1$, then $(a+b)^p \le 2^{p-1}(a^p + b^p)$. Assume the following conditions hold,

$$\begin{aligned} \sup_k E\left\{\left|h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right|^2 \Big|\, \mathcal{F}_k\right\} &< \infty \quad \text{a.s.}, \\ \sup_k E\left\{\left|h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right|^2 \Big|\, \mathcal{F}_k\right\} &< \infty \quad \text{a.s.}, \\ \sup_k E\left\{|S_{\theta_k}(\omega_{k+1})|^2 \Big|\, \mathcal{F}_k\right\} &< \infty \quad \text{a.s.} \end{aligned}$$

Then,

$$
\begin{aligned}
& \sum_{k=0}^{\infty} \alpha_{k+1}^2 E\left\{ \left| W_{k+1}^T \bar{h}(\theta_k) \right|^2 \middle| \mathcal{F}_k \right\} \\
\leq\ & C \sum_{k=0}^{\infty} \alpha_{k+1}^2 E\left\{ \left| W_{k+1} \right|^2 \middle| \mathcal{F}_k \right\} \\
\leq\ & C \sum_{k=0}^{\infty} \alpha_{k+1}^2 \sup_k E\left\{ \left| W_{k+1} \right|^2 \middle| \mathcal{F}_k \right\} \\
<\ & \infty
\end{aligned}
$$

almost surely, where the constant C is the bound on $|\bar{h}(\theta)|^2$. The MCT [4, Theorem 2.17] then states $\sum_{k=0}^{\infty} \alpha_{k+1} W_{k+1}^T \bar{h}(\theta_k)$ exists almost surely.

We summarise the above analysis of $\liminf_k \left|\bar{h}(\theta_k)\right|$ in the following lemma.

Lemma 7 *Consider the recursion for θ_k re-written as $\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}$, where $H_{k+1} = -\Gamma(b_k)\bar{h}(\theta_k) + W_{k+1}$, and the 'noise' term W_{k+1} is given by*

$$
\begin{aligned}
W_{k+1} &= -\Gamma(b_k)\left(h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k)\right) - \Gamma(b_k)\left(h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k)\right) \\
&\quad + \Gamma(b_k) S_{\theta_k}(\omega_{k+1}) b_k,
\end{aligned} \tag{9}
$$

and satisfies $E_k(W_{k+1}) = 0$. Let the following assumptions hold,
(a) $\bar{h}(\theta)$ $(= \nabla J(\theta))$ satisfies $\sup_\theta |\bar{h}(\theta)| < \infty$ and, for some constant C and all θ, θ',

$$
|\bar{h}(\theta) - \bar{h}(\theta')| < C\,|\theta - \theta'|\,;
$$

(b)

$$
\begin{aligned}
\sup_k E\left(|H_k|^2\right) &< \infty, \\
\sup_k E\left\{ \left| h_{1,\theta_k}(\omega_{k+1}) - \bar{h}_1(\theta_k) \right|^2 \middle| \mathcal{F}_k \right\} &< \infty \qquad a.s., \\
\sup_k E\left\{ \left| h_{2,\theta_k}(\omega_{k+1}) - \bar{h}_2(\theta_k) \right|^2 \middle| \mathcal{F}_k \right\} &< \infty \qquad a.s., \\
\sup_k E\left\{ \left| S_{\theta_k}(\omega_{k+1}) \right|^2 \middle| \mathcal{F}_k \right\} &< \infty \qquad a.s.;
\end{aligned}
$$

(c) $\sup_k |b_k| < \infty$ almost surely.
Then, if J is bounded below,

$$
\liminf_k |\nabla J(\theta_k)| = 0.
$$

Remark 8 *Lemma 7 could have been stated in a slightly more general and concise way by omitting the declaration of W_k in (9) entirely, and replacing assumption (b) with*
(b') $\sum_{k=0}^{\infty} \alpha_{k+1} W_{k+1}^T \bar{h}(\theta_k)$ exists almost surely.

2 A Variant of the Main Algorithm

In this section, we present a variant of the main algorithm (1)-(3) that is more suitable for the OTP application. As we show, the convergence results for the main algorithm carry over easily.

Let

$$
F = [F_1, \cdots, F_d]^T : R^d \to R^d
$$

be bounded mapping with bounded partial derivatives, i.e.,

$$\sup_{\theta \in R^d} |F(\theta)| \quad < \quad \infty, \tag{10}$$

$$\sup_{\theta \in R^d} |\nabla F(\theta)| \quad = \quad |[\nabla F_1(\theta), \cdots, \nabla F_d(\theta)]| < \infty. \tag{11}$$

(Note that $\nabla F_i = [\frac{\partial F_i}{\partial \theta_1}, \cdots, \frac{\partial F_i}{\partial \theta_d}]^T$.) As detailed in Section 3 below, introducing the mapping F is important for the OTP application. It follows that

$$\begin{aligned} \nabla(J \circ F)(\theta) &= \nabla F(\theta) \nabla J(F(\theta)) \\ &= \nabla F(\theta) \left(\bar{h}_1(F(\theta)) + \bar{h}_2(F(\theta))\right) \\ &= \nabla F(\theta) E_{\omega \sim P_{F(\theta)}} \left\{ h_{1,F(\theta)}(\omega) + h_{2,F(\theta)}(\omega) \right\}. \end{aligned}$$

Drawing parallels with algorithm (1)-(3) (and the discussion) in Section 1 suggests the following algorithm to minimise $J \circ F$,

$$\begin{aligned} \theta_{k+1} &= \theta_k - \alpha_{k+1} \Gamma(b_k) \nabla F(\theta_k) \\ &\quad \times \left(h_{1,F(\theta_k)}(\omega_{k+1}) + h_{2,F(\theta_k)}(\omega_{k+1}) - S_{F(\theta_k)}(\omega_{k+1}) b_k \right), && (12) \\ b_{k+1} &= b_k - \beta_{k+1} \left(S^2_{\tilde{\theta}_k}(\omega_{k+1}) b_k - S_{\tilde{\theta}_k}(\omega_{k+1}) h_{1,\tilde{\theta}_k}(\omega_{k+1}) \right), && (13) \\ \omega_{k+1} &\sim P_{\tilde{\theta}_k}, && (14) \\ \tilde{\theta}_k &= F(\theta_k), \qquad k \geq 0. && (15) \end{aligned}$$

This is the main algorithm of the report.

Let the range of the function F be denoted by range(F). Assuming F has bounded partial derivatives we have,

$$|F(\theta) - F(\theta')| \leq C \, |\theta - \theta'|$$

for some constant C. Since $\tilde{\theta}_{k+1} = \tilde{\theta}_k + \alpha_{k+1} \frac{(\tilde{\theta}_{k+1} - \tilde{\theta}_k)}{\alpha_{k+1}}$, we also have

$$\left| \frac{\tilde{\theta}_{k+1} - \tilde{\theta}_k}{\alpha_{k+1}} \right| \leq \frac{C}{\alpha_{k+1}} \, |\theta_{k+1} - \theta_k| = C |H_{k+1}|,$$

which implies that it is sufficient to verify Assumption 3 with

$$H_{k+1} = -\Gamma(b_k) \nabla F(\theta_k) \left(h_{1,F(\theta_k)}(\omega_{k+1}) + h_{2,F(\theta_k)}(\omega_{k+1}) - S_{F(\theta_k)}(\omega_{k+1}) b_k \right). \tag{16}$$

The following corollary to Lemma 2 is straightforward.

Corollary 9 *Consider Algorithm (12)-(15). Let Assumption 1 be satisfied, Assumption 2a for all $\theta \in$range(F), 2b for all $\theta, \theta' \in$range(F) and 2c with θ_k replaced by $\tilde{\theta}_k$ defined in (15). Let Assumption 3 be satisfied with H_{k+1} given in (16) and Assumption 4 for all $b \in R^d, \theta \in$range(F). Then, almost surely,* $\sup_k |b_k| < \infty$ *and* $\lim_k \left| b_k - \overline{S^2}(\tilde{\theta}_k)^{-1} \overline{S \times h_1}(\tilde{\theta}_k) \right| = 0$.

We now state the equivalence of Lemma 7 for the modified algorithm (12)-(15) in the following corollary.

Corollary 10 *Consider the recursion for θ_k above re-written as $\theta_{k+1} = \theta_k + \alpha_{k+1} H_{k+1}$ where $H_{k+1} = -\Gamma(b_k) \nabla F(\theta_k) \bar{h}(\theta_k) + W_{k+1}$, and noise term*

$$\begin{aligned} W_{k+1} &= -\Gamma(b_k) \nabla F(\theta_k) \left(h_{1,F(\theta_k)}(\omega_{k+1}) - \bar{h}_1(F(\theta_k)) \right) \\ &\quad -\Gamma(b_k) \nabla F(\theta_k) \left(h_{2,F(\theta_k)}(\omega_{k+1}) - \bar{h}_2(F(\theta_k)) \right) \\ &\quad +\Gamma(b_k) \nabla F(\theta_k) S_{F(\theta_k)}(\omega_{k+1}) b_k \end{aligned}$$

satisfies $E_k(W_{k+1}) = 0$. *Assume*
(a) $\nabla F(\theta)\bar{h}(F(\theta))$ *(*$= \nabla(J \circ F)(\theta)$*) satisfies* $\sup_\theta |\nabla(J \circ F)(\theta)| < \infty$ *and, for some constant* C *and all* θ, θ',

$$|\nabla(J \circ F)(\theta) - \nabla(J \circ F)(\theta')| < C\,|\theta - \theta'|\,;$$

(b) that

$$\begin{aligned}
\sup_k E\left(|H_k|^2\right) &< \infty, \\
\sup_k E\left\{ \left|\nabla F(\theta_k)\left(h_{1,F(\theta_k)}(\omega_{k+1}) - \bar{h}_1(F(\theta_k))\right)\right|^2 \Big| \mathcal{F}_k \right\} &< \infty, \quad a.s., \qquad (17)\\
\sup_k E\left\{ \left|\nabla F(\theta_k)\left(h_{2,F(\theta_k)}(\omega_{k+1}) - \bar{h}_2(F(\theta_k))\right)\right|^2 \Big| \mathcal{F}_k \right\} &< \infty, \quad a.s., \qquad (18)\\
\sup_k E\left\{ \left|\nabla F(\theta_k) S_{F(\theta_k)}(\omega_{k+1})\right|^2 \Big| \mathcal{F}_k \right\} &< \infty, \quad a.s.; \qquad (19)
\end{aligned}$$

(c) $\sup_k |b_k| < \infty$ *almost surely.*
Then, if $J \circ F$ *is bounded below,*

$$\liminf_k |\nabla(J \circ F)(\theta_k)| = 0, \qquad \textit{almost surely.}$$

Remark 11 *In Corollary 10, if we assume that function* F *has bounded first and second order derivatives*[2]*, we have,*

$$\begin{aligned}
|F(\theta) - F(\theta')| &\leq C_1\,|\theta - \theta'|\,, \\
|\nabla F(\theta) - \nabla F(\theta')| &\leq C_2\,|\theta - \theta'|\,,
\end{aligned}$$

and the Lipschitz condition on $\nabla(J \circ F)$ *is satisfied provided* $|\bar{h}(F(\theta))|$ *is bounded and* ∇J *Lipschitz when its domain is restricted to range(*F*). This follows since,*

$$\begin{aligned}
& |\nabla F(\theta)\bar{h}(F(\theta)) - \nabla F(\theta')\bar{h}(F(\theta'))| \\
\leq \; & |\bar{h}(F(\theta))||\nabla F(\theta) - \nabla F(\theta')| + |\nabla F(\theta')||\bar{h}(F(\theta)) - \bar{h}(F(\theta'))|.
\end{aligned}$$

Remark 12 *When* F *is bounded, to verify (17)-(19) in Corollary 10, it is sufficient to verify*

$$\begin{aligned}
\sup_{\tilde{\theta} \in range(F)} E_{\omega \sim P_{\tilde{\theta}}} \left\{ \left| h_{1,\tilde{\theta}}(\omega) - \bar{h}_1(\tilde{\theta}) \right|^2 \right\} &< \infty, \\
\sup_{\tilde{\theta} \in range(F)} E_{\omega \sim P_{\tilde{\theta}}} \left\{ \left| h_{2,\tilde{\theta}}(\omega) - \bar{h}_2(\tilde{\theta}) \right|^2 \right\} &< \infty, \\
\sup_{\tilde{\theta} \in range(F)} E_{\omega \sim P_{\tilde{\theta}}} \left\{ \left| S_{\tilde{\theta}}(\omega) \right|^2 \right\} &< \infty.
\end{aligned}$$

3 Trajectory Planning Problem

A more detailed account of the OTP problem is contained in [7]. Below, we only present the relevant material from [7] for the convergence to be studied and the report to be self-contained.

At time n, let $X_n \in R^{d_x}$ and $Y_n \in R^{d_y}$ be the random vectors that model the d_x-dimensional state and its d_y-dimensional observation respectively. Suppose that an action $A_n \in R^{d_a}$ is applied at time n. The state $\{X_n\}_{n \geq 0}$ is an unobserved Markov process with initial distribution and transition law given by

$$X_0 \sim \pi_0, \quad X_{n+1} \sim p\left(\cdot|X_n\right), \qquad (20)$$

[2] $\sup_\theta \left| \frac{\partial^2 F_i(\theta)}{\partial \theta_k \partial \theta_j} \right| < \infty$ for each i, k, j.

respectively. The observation process $\{Y_n\}_{n\geq 0}$ is generated according to the state and action dependent probability density

$$Y_n \sim q\left(\cdot|X_n, A_n\right). \tag{21}$$

In OTP, we wish to track a maneuvering target for N epochs. At epoch n, X_n denotes the state of the target, A_n the position of the observer and Y_n the partial observation of the target state, i.e., $Y_n = g(X_n, A_n, V_n)$, where V_n is measurement noise. The aim of OTP is to adaptively maneuver the observer to optimise the tracking performance the target.

Given the sequence of actions $a_{1:n} := \{a_1, ..., a_n\}$ and measurements $y_{1:n} := \{y_1, ..., y_n\}$, the *filtering density* at time n is denoted by π_n (or $\pi_n^{(y_{1:n},a_{1:n})}$ to emphasise the dependence on $y_{1:n}$, $a_{1:n}$) and satisfies the *Bayes* recursion

$$\pi_n(x) = \frac{q\left(y_n|x, a_n\right)\int p(x|x')\pi_{n-1}(x')dx'}{\int\int q\left(y_n|x, a_n\right)p(x|x')\pi_{n-1}(x')dx'dx}. \tag{22}$$

The initial state density π_0 is fixed. The full posterior, i.e., the density of $X_{0:n}$ given $Y_{1:n}$ and $A_{1:n}$ is,

$$\pi_{0:n}(x_{0:n}) = \frac{\left(\Pi_{i=1}^{n} q\left(Y_i|x_i, A_i\right)p\left(x_i|x_{i-1}\right)\right)\pi_0(x_0)}{\int\left(\Pi_{i=1}^{n} q\left(Y_i|x_i, A_i\right)p\left(x_i|x_{i-1}\right)\right)\pi_0(x_0)dx_{0:n}}.$$

The simulation-based algorithm proposed in [7] for sensor scheduling requires both the marginal π_n and the full posterior $\pi_{0:n}$ for all N epochs, i.e., for $1 \leq n \leq N$. In [7] we proposed to approximate these densities using a mixture of Dirac delta-masses,

$$\hat{\pi}_{0:n}(x_{0:n}) := \sum_{j=1}^{L} w_n^{(j)}\delta_{X_{0:n}^{(j)}}(x_{0:n}), \tag{23}$$

where $\delta_{X_{0:n}^{(j)}}$ denotes the Dirac delta-mass located at $X_{0:n}^{(j)}$ and the *importance weights* $\{w_n^{(j)}\}_{j=1}^{L}$ are non-negative scalars that sum to one. The approximation to π_n, namely $\hat{\pi}_n$, follows by marginalising $\hat{\pi}_{0:n}$. There are a number of ways to define such a point-mass approximation. For example, the simplest scheme would be to sample L independent state trajectory realisations $\left\{X_{0:n}^{(j)}\right\}_{j=1}^{L}$ from $\left(\Pi_{i=1}^{n} p\left(x_i|x_{i-1}\right)\right)\pi_0(x_0)$. The importance weights would then be

$$w_n^{(j)} := \frac{\Pi_{i=1}^{n} q\left(Y_i|X_i^{(j)}, A_i\right)}{\sum_{j=1}^{L}\Pi_{i=1}^{n} q\left(Y_i|X_i^{(j)}, A_i\right)}. \tag{24}$$

For any integrable function h, $\int h(x_{0:n})\hat{\pi}_{0:n}(x_{0:n})dx_{0:n}$ converges to $\int h(x_{0:n})\pi_{0:n}(x_{0:n})dx_{0:n}$ as $L \to \infty$ (see [3, Ch. 2] for a precise statement of the mode of convergence). Practically though, we would prefer a small sample size L and this simple scheme of sampling from the state transition model can result in the majority of the importance weights $w_n^{(j)}$ being very small. There are number of remedies proposed for this in the Sequential Monte Carlo literature [3, Ch. 1.3.2]. For example, the importance sampling step can be designed to minimise the conditional variance of the importance weights by sampling $\left\{X_{0:n}^{(j)}\right\}_{j=1}^{L}$ from a Markov transition density that takes the observations into account. We emphasise that standard techniques from the Sequential Monte Carlo literature can be adopted in constructing an approximation of the form (23) to the full posterior, but we do not study this issue in detail here. To simplify the presentation, we will only focus on the simple scheme of sampling from the state transition model. To emphasise the dependence of $\hat{\pi}_{0:n}$ on the realisation of observations $Y_{1:n}$ and the sequence of actions $A_{1:n}$, we use the notation $\hat{\pi}_{0:n}^{(Y_{1:n},A_{1:n})}$. We often drop the sequence of observations and\or actions from $\hat{\pi}_{0:n}$ in order to unclutter the expressions. **The reader is reminded though that $\hat{\pi}_{0:n}$ should always be regarded as a function of $(Y_{1:n}, A_{1:n})$.**

Consider a suitable *test* function $\psi : \left(R^{d_x}\right)^{N+1} \to R$. (For example, ψ could pick out a component of interest of the state vector.) In [7], it was proposed to solve the following sensor scheduling problem,

$$\min_{A_{1:N} \in \Theta_A} J(A_{1:N}) = E_{A_{1:N}} \left\{ \left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle^2 \right\}, \tag{25}$$

where $\Theta_A \subset \left(R^{d_a}\right)^N$ is an open set.[3] The expectation operator is to be understood in the following sense. For an integrable function $h : \left(R^{d_x}\right)^{N+1} \times \left(R^{d_a}\right)^N \times \left(R^{d_y}\right)^N \to R$,

$$\begin{aligned} & E_{A_{1:N}} \left\{ h(X_{0:N}, A_{1:N}, Y_{1:N}) \right\} \\ & := \int h(x_{0:N}, A_{1:N}, y_{1:N}) \Pi_{i=1}^N q\left(y_i | x_i, A_i\right) p\left(x_i | x_{i-1}\right) \pi_0(x_0) dx_{0:N} dy_{1:N}. \end{aligned} \tag{26}$$

3.1 Bearings-only Tracking

We do not need to specify the target model explicitly. Our only concern is that we can sample from the model. Maneuvering targets are often modelled such as a *jump Markov linear system* (JMLS) [2]. The state of the target is comprised of continuous and discrete valued variables, i.e.,

$$X_n = \left[r_{x,n}, v_{x,n}, r_{y,n}, v_{y,n}, \xi_n\right]^T \in R^4 \times \Xi,$$

where $(r_{x,n}, r_{y,n})$ denotes the target's (Cartesian) coordinates at time n, $(v_{x,n}, v_{y,n})$ denotes the target's velocity in the x and y directions, and ξ_n denotes the mode of the target, which belongs to a finite set Ξ. The target switches discontinuously, as indicated by ξ_n, between constant velocity maneuvers.

Let the observer model be of the form

$$A_{1:N} = F(U_{1:N}),$$

where we exert control on the observer positions $A_{1:N}$ through the variables $U_{1:N}$. For instance, the accelerations of the observer could be determined from the input $U_{1:N}$, which will in turn determine the observer trajectory. The convergence results of Propositions 14 and 15 below do not depend on the specific form of F but only that this function is sufficiently regular. We now give an example of F which was adopted in the numerical section of [7].

Example 13 *Let the state of the observer be*

$$X_n^o = \left[r_{x,n}^o, v_{x,n}^o, r_{y,n}^o, v_{y,n}^o\right]^T, \tag{27}$$

with

$$A_n = \left[r_{x,n}^o, r_{y,n}^o\right]^T. \tag{28}$$

Assume the following kinematic model for the evolution of the state,

$$X_{n+1}^o = \underbrace{\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{=:G} X_n^o + \underbrace{\begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix}}_{=:H} \times C \times atan(U_{n+1}) \tag{29}$$

where the initial state X_0^o is fixed, T is the sampling interval, and $U_{n+1} \in R^2$ determines the acceleration in the x and y directions. We have included the function atan and the positive diagonal matrix C. The function atan and its first two derivatives are bounded. Also, atan is linear about zero and makes a nice choice of

[3] Problem (25) corresponds to [7, eqn. (16)] with a discount factor $\lambda = 0$. The same analysis applies for $\lambda \in (0, 1]$.

saturating function for the acceleration; naturally the acceleration cannot be unbounded. The matrix C alters the saturation behaviour of the acceleration. The observer trajectory is completely determined once X_0^o and $U_{1:N}$ are given,

$$A_n = \begin{bmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \end{bmatrix} \times \left(G^n X_0^o + \sum_{i=1}^{n} G^{n-i} HC\,atan(U_i) \right). \tag{30}$$

The function F is now implicitly defined.

In the bearings-only model, the observation process $\{Y_n\}_{n\geq 0}$ $(\subset R)$ is generated according to

$$Y_n = \operatorname{atan}\left(\frac{r_{x,n} - A_n(1)}{r_{y,n} - A_n(2)} \right) + V_n, \tag{31}$$

where $V_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_Y^2)$. In our simulation-based framework, we require that the observation process density is known and is differentiable w.r.t. A_n. The bearings-only case is one such example. To present the convergence results of Proposition 14 and 15 below, we will assume that the x and y position of the target corresponds to the first and third component of the state descriptor X_n,

$$X_n = \left[r_{x,n}, \cdot, r_{y,n}, \cdots \right]^T, \tag{32}$$

which is usual convention in the literature.

Before stating the two timescale SA algorithm for OTP, we first define the following functions. The R^2 vector-valued function called the *score* is defined to be,

$$\begin{aligned} S(y,x,a)^T &= \left[\nabla_{a(1)} q(y|x,a), \qquad \nabla_{a(2)} q(y|x,a) \right] \times q(y|x,a)^{-1} \\ &= \left[S_1(y,x,a), \qquad S_2(y,x,a) \right] \\ &= -\sigma_Y^{-2} \frac{y - \operatorname{atan}\left(\frac{x(1)-a(1)}{x(3)-a(2)} \right)}{1 + \left(\frac{x(1)-a(1)}{x(3)-a(2)} \right)^2} \left[\frac{1}{x(3)-a(2)}, \qquad -\frac{x(1)-a(1)}{\left(x(3)-a(2) \right)^2} \right]. \end{aligned} \tag{33}$$

For each $A_{1:N} \in R^{2N}$, define the diagonal matrix-valued function $S_{A_{1:N}} : (R^{d_x})^{N+1} \times (R)^N \to R^{2N} \times R^{2N}$ as follows,

$$S_{A_{1:N}}(X_{0:N}, Y_{1:N}) = \operatorname{diag}\left(\left[[S_j(Y_i, X_i, A_i)]_{j=1,2} \right]_{i=1,\dots,N} \right). \tag{34}$$

Note that $S_{A_{1:N}}(X_{0:N}, Y_{1:N})$ is just the score stacked as a vector, and then converted in to a diagonal matrix. For each $A_{1:N} \in R^{2N}$, let functions $h_{1,A_{1:N}}, h_{2,A_{1:N}} : (R^{d_x})^{N+1} \times (R)^N \to (R^2)^N$ be given as follows,

$$\begin{aligned} &h_{1,A_{1:N}}(X_{0:N}, Y_{1:N})) = \left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle^2 \right) S_{A_{1:N}}(X_{0:N}, Y_{1:N}) \mathbf{1}, \\ &h_{2,A_{1:N}}(X_{0:N}, Y_{1:N}) \\ &= \left[\left[\begin{array}{c} \left\langle \psi^2 S_j(Y_i, \cdot, A_i), \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \\ - \left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \left\langle S_j(Y_i, \cdot, A_i), \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \\ -2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \left\langle \psi S_j(Y_i, \cdot, A_i), \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \\ +2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle^2 \left\langle S_j(Y_i, \cdot, A_i), \hat{\pi}_{0:N}^{(Y_{1:N}, A_{1:N})} \right\rangle \end{array} \right]_{j=1,2} \right]_{i=1,\dots,N}, \end{aligned}$$

where $\mathbf{1}$ is the vector with elements 1. (*Notation:* For scalars $a_{j,i}$, $j = 1,\dots,m$, $i = 1,\dots,n$, $\left[[a_{j,i}]_{j=1,\dots,m} \right]_{i=1,\dots,n}$ denotes the stacked vector $[a_{1,1},\dots,a_{m,1},\dots,a_{1,n},\dots,a_{m,n}]^T$; see the declaration of notation in the Introduction.) One may verify that the gradient of the performance criterion satisfies[4]

$$\nabla J(A_{1:N}) = E_{A_{1:N}} \left\{ h_{1,A_{1:N}}(X_{0:N},Y_{1:N})) + h_{2,A_{1:N}}(X_{0:N},Y_{1:N}) \right\}. \tag{35}$$

The aim is to study the convergence of the following two timescale SA algorithm for OTP that was presented in [7]:

$$\begin{aligned} U_{1:N,k+1} &= U_{1:N,k} - \alpha_{k+1}\Gamma(b_k)\nabla F(U_{1:N,k}) \\ &\quad \times \big(h_{1,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1}) \\ &\quad + h_{2,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1}) \\ &\quad - S_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})b_k\big), \end{aligned} \tag{36}$$

$$\begin{aligned} b_{k+1} &= b_k - \beta_{k+1}S^2_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})b_k \\ &\quad + \beta_{k+1}S_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1}) \\ &\quad \times h_{1,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})), \end{aligned} \tag{37}$$

$$(X_{0:N,k+1},Y_{1:N,k+1}) \sim P_{A_{1:N,k}}, \tag{38}$$

$$A_{1:N,k} = F(U_{1:N,k}), \qquad k \geq 0, \tag{39}$$

where $A_{1:N,k}, U_{1:N,k} \in R^{2N}$. If we set $\theta_k = U_{1:N,k}$, $\tilde{\theta}_k = A_{1:N,k}$ and $\omega_{k+1} = (X_{0:N,k+1},Y_{1:N,k+1})$ then, the correspondence between this algorithm and Algorithm (12)-(15) is obvious.

Define

$$\overline{S^2}(A_{1:N}) = E_{A_{1:N}} \left\{ S^2_{A_{1:N}}(X_{0:N},Y_{1:N}) \right\}, \tag{40}$$

$$\overline{S \times h_1}(A_{1:N}) = E_{A_{1:N}} \left\{ S_{A_{1:N}}(X_{0:N},Y_{1:N}) h_{1,A_{1:N}}(X_{0:N},Y_{1:N}) \right\}. \tag{41}$$

For the bearings-only observation model, we have the following result concerning the convergence of the fast and slow timescale.

Proposition 14 *Consider Algorithm (36)-(39) for the bearings-only observation model corrupted by Gaussian additive noise (31). Suppose that the following assumptions hold,*

$$\sup_{A_{1:N}\in range(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left| \frac{1}{X_i(3)-A_i(2)} \right|^p \right\} < \infty, \quad 1 \leq i \leq N,\ p > 0, \tag{42}$$

$$\sup_{A_{1:N}\in range(F)} \max_l \left| \frac{1}{X_i^{(l)}(3)-A_i(2)} \right| < \infty, \quad 1 \leq i \leq N, \tag{43}$$

$$\inf_{1\leq i\leq N} \inf_{A_{1:N}\in range(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \sigma_Y^{-2} \frac{\frac{1}{(X_i(3)-A_i(2))^2}}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2} \right\} > 0, \tag{44}$$

$$\inf_{1\leq i\leq N} \inf_{A_{1:N}\in range(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \sigma_Y^{-2} \frac{\left[\frac{X_i(1)-A_i(1)}{(X_i(3)-A_i(2))^2}\right]^2}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2} \right\} > 0. \tag{45}$$

[4]Note that

$$\begin{aligned} & \frac{\partial}{\partial a_i(j)} \left\langle \psi, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})} \right\rangle, \\ = & \left\langle \psi(\cdot)S_j(y_i,\cdot,a_i), \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})} \right\rangle \left\langle S_j(y_i,\cdot,a_i), \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})} \right\rangle. \end{aligned}$$

Then, almost surely, $\sup_k |b_k| < \infty$ *and* $\lim_k \left| b_k - \overline{S^2}(A_{1:N,k})^{-1}\overline{S \times h_1}(A_{1:N,k}) \right| = 0$. *Furthermore, if F has bounded second order derivatives then, almost surely,*

$$\liminf_k |\nabla(J \circ F)(U_{1:N,k})| = 0.$$

The proof of Proposition 14 appears in Section 4. Condition (43) relates to the samples used to approximate the posterior density (23)-(24). Also, the first and third component of the target state is its x and y coordinate respectively. Note that the proposition does not limit the specific form of function F that relates inputs $U_{1:N}$ to actions $A_{1:N}$. It only requires F to be sufficiently regular as specified by the last assumption concerning bounded second order derivatives. For F defined implicitly by (30), this assumption is satisfied.

The next result gives the conditions under which assumptions (42)-(45) hold. The result basically says that if the support of $X_{0:N}$ and that the range of F do not intersect, then the assumptions hold and we have the desired convergence of two timescale SA for OTP. It is interesting to note that the scenario where this is true is that studied by all previous works on OTP for bearings-only observations [7]. Thus, the conditions of Proposition 14 do not appear to be too restrictive for the application.

Proposition 15 *Write the mapping* $F : R^{2N} \to R^{2N}$ *as* $F = [F_{1,1}, F_{1,2}, \cdots, F_{N,1}, F_{N,2}]^T$. *(Note that* $A_i(j) = F_{i,j}(U_{1:N})$.*) Suppose that the density of* $X_{0:N}$, $f(x_{0:N})$, *has a compact support* $\mathcal{K}_f \subset R^{4(N+1)}$. *Furthermore, suppose that for each* $1 \le i \le N$, *the compact set* $\mathcal{K}_{f,i} := \{ x_i(3) | \, x_{0:N} \in \mathcal{K}_f\}$ *does not intersect with the closure of the set range*$(F_{i,2})$, *i.e., there exists a compact set* $\mathcal{K}_{A,i}$ *such that range*$(F_{i,2}) \subset \mathcal{K}_{A,i}$, *and* $\mathcal{K}_{f,i} \cap \mathcal{K}_{A,i} = \varnothing$. *Then, conditions (42)-(45) are satisfied.*

Proof. Note that $\inf_{X_i(3)\in\mathcal{K}_{f,i}, A_i(2)\in\mathcal{K}_{A,i}} |X_i(3) - A_i(2)| > 0$ since we a minimising a continuous function over the compact set $\mathcal{K}_{f,i} \times \mathcal{K}_{A,i}$; at least one solution from this compact set exists. Condition (42) now follows since for any $A_{1:N}$ $\in$range(F),

$$\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left| \frac{1}{X_i(3) - A_i(2)} \right|^p \right\} \\
= \; & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ I_{\mathcal{K}_{f,i}}(X_i(3)) \left| \frac{1}{X_i(3) - A_i(2)} \right|^p \right\} \\
\le \; & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left(\sup_{X_i(3)\in\mathcal{K}_{f,i}, A_i(2)\in\mathcal{K}_{A,i}} \frac{1}{|X_i(3) - A_i(2)|} \right)^p \right\} \\
\le \; & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left(\frac{1}{\inf_{X_i(3)\in\mathcal{K}_{f,i}, A_i(2)\in\mathcal{K}_{A,i}} |X_i(3) - A_i(2)|} \right)^p \right\}.
\end{aligned}$$

A similar argument establishes (43). As for (45), note that

$$\begin{aligned}
& \inf_{X_{0:N}\in\mathcal{K}_f, A_{1:N}\in\text{range}(F)} \frac{\frac{1}{(X_i(3)-A_i(2))^4}}{\left[1 + \left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)} \right)^2 \right]^2} \\
\ge \; & \frac{\frac{1}{\left(\sup_{X_i(3)\in\mathcal{K}_{f,i}, A_i(2)\in\mathcal{K}_{A,i}} |X_i(3)-A_i(2)| \right)^4}}{\left[1 + \left(\frac{\sup_{X_{0:N}\in\mathcal{K}_f, A_{1:N}\in\text{range}(F)} |X_i(1)-A_i(1)|}{\inf_{X_i(3)\in\mathcal{K}_{f,i}, A_i(2)\in\mathcal{K}_{A,i}} |X_i(3)-A_i(2)|} \right)^2 \right]^2} = b(>0).
\end{aligned}$$

Thus,

$$
\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\frac{\left[\frac{X_i(1)-A_i(1)}{(X_i(3)-A_i(2))^2}\right]^2}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2}\right\} \\
= \;& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{I_{\mathcal{K}_f}(X_{0:N})\frac{\left[\frac{X_i(1)-A_i(1)}{(X_i(3)-A_i(2))^2}\right]^2}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2}\right\} \\
\geq \;& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\begin{array}{c} I_{\mathcal{K}_f}(X_{0:N})\left(X_i(1)-A_i(1)\right)^2 \\ \times \inf_{X_{0:N}\in\mathcal{K}_f, A_{1:N}\in \mathrm{range}(F)} \frac{\frac{1}{(X_i(3)-A_i(2))^4}}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2}\end{array}\right\} \\
\geq \;& bE_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(X_i(1)-A_i(1)\right)^2\right\} \\
\geq \;& b\times var\left\{X_i(1)\right\},
\end{aligned}
$$

where in the last line, the density of $X_{0:N}$ is independent of the sequence of actions $A_{1:N}$ and hence we write $var\{\cdot\}$ omitting reference to the actions. A similar argument establishes (44). ■

4 Proof of Proposition 14

4.1 Verifying the assumptions for convergence of the fast timescale

Proposition 14 stipulates a set of conditions under which we have convergence of the fast and slow timescale. We begin by showing how satisfying (42)-(45) would imply the assumptions of Corollary 9 are satisfied and hence proving convergence of the fast timescale in Proposition 14.

4.1.1 Assumptions 2a and 2c

We will show that a sufficient condition for Assumption 2a and 2c is

$$
\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^p\right\} < \infty, \qquad 1\leq i\leq N, \quad p>0,
$$

i.e., (42) of Proposition 14.

Recall the definition of $h_{1,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1}))$,

$$
\begin{aligned}
& h_{1,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})) \\
= \;& \left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})}\right\rangle^2\right) S_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})\mathbf{1}.
\end{aligned}
$$

Note that the term $\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})}\right\rangle^2\right)$ is bounded by

$$
2\max_i \psi^2(X_{0:N}^{(i)}).
$$

Using this fact,

$$\begin{aligned}
& \left|\overline{S \times h_1}(A_{1:N})\right| \\
= & \left|E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_{A_{1:N}}(X_{0:N},Y_{1:N})h_{1,A_{1:N}}(X_{0:N},Y_{1:N})\right\}\right| \\
\leq & C \times \left|E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S^2_{A_{1:N}}(X_{0:N},Y_{1:N})\mathbf{1}\right\}\right| \\
\leq & C \times |\mathbf{1}| \times \left|\overline{S^2}(A_{1:N})\right|,
\end{aligned}$$

where C is the bound $2\max_i \psi^2(X^{(i)}_{0:N})$. Thus we only have to verify the boundedness of $\left|\overline{S^2}(A_{1:N})\right|$ to verify Assumption 2a.

As for Assumption 2c, we have similarly

$$\begin{aligned}
& \left|S_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})h_{1,A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})\right| \\
\leq & C \times |\mathbf{1}| \times \left|S^2_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})\right|.
\end{aligned}$$

Thus, Assumption 2c is satisfied provided

$$\begin{aligned}
& \sup_k E_{(X_{0:N,k+1},Y_{1:N,k+1})\sim P_{A_{1:N,k}}}\left\{\left|S^2_{A_{1:N,k}}(X_{0:N,k+1},Y_{1:N,k+1})\right|^p\right\} \\
\leq & \sup_{A_{1:N}\in \mathrm{range}(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|S^2_{A_{1:N}}(X_{0:N},Y_{1:N})\right|^p\right\} \\
< & C_p.
\end{aligned}$$

As $S^2_{A_{1:N}}(X_{0:N},Y_{1:N})$ is a diagonal matrix, it is sufficient to verify

$$\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|S^2_j(Y_i,X_i,A_i)\right|^p\right\} < C_{p,j,i}$$

for each $p > 0$, $j \in \{1,2\}$, $i = 1,\ldots,N$. In fact, verifying this will also simultaneously verify Assumption 2a.

$$\begin{aligned}
|S_1(Y_i,X_i,A_i)|^p &= \sigma_Y^{-2p} \times \left|\frac{Y_i - \mathrm{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)}{1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2}\right|^p \left|\frac{1}{X_i(3)-A_i(2)}\right|^p \\
&\leq \sigma_Y^{-2p}\left|Y_i - \mathrm{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\right|^p \left|\frac{1}{X_i(3)-A_i(2)}\right|^p
\end{aligned}$$

Thus, $E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\sigma_Y^{-2p}\left|Y_i - \mathrm{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\right|^p\right\} \leq C_p$ when p is even where constant C_p follows from the central moment of a Gaussian. When p is odd, $E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\sigma_Y^{-2p}\left|Y_i - \mathrm{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\right|^p\right\} \leq \sigma_Y^{-2p} + \sigma_Y^2 C_{p+1}$, which follows from the inequality $|x|^p \leq 1 + |x|^{p+1}$. Thus $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\{|S_1(Y_i,X_i,A_i)|^p\} < \infty$ for all i and p provided

$$\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^p\right\} < \infty, \qquad 1 \leq i \leq N, \quad p > 0. \tag{46}$$

A similar argument shows $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\{|S_2(Y_i,X_i,A_i)|^p\} < \infty$ provided (46) holds.

4.1.2 Assumption 2b

We need to verify Lipschitz continuity of $\overline{S^2}(A_{1:N})$. Since

$$\begin{aligned}\left|\overline{S^2}(A_{1:N})-\overline{S^2}(A'_{1:N})\right| \quad \le \quad & \max_{j,i}\Big|E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\} \\ & -E_{(X_{0:N},Y_{1:N})\sim P_{A'_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\}\Big|,\end{aligned}$$

it is sufficient to verify, for each i,j, Lipschitz continuity of the function $A_{1:N}\to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\}$. By the Mean Value Theorem, the mapping $A_{1:N}\to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\}$ is Lipschitz continuous if its partial derivatives are bounded. We will show that a sufficient condition for the Lipschitz continuity of $\overline{S^2}(A_{1:N})$ is

$$\sup_{A_{1:N}\in\text{range}(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^3\right\}<\infty, \qquad 1\le i\le N, \tag{47}$$

which is implied by (42) of Proposition 14.

Likewise, since

$$\begin{aligned}&\left|\overline{S\times h_1}(A_{1:N})-\overline{S\times h_1}(A'_{1:N})\right| \\ \le\quad & \max_{j,i}\sqrt{2N}\,\Big|E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle\psi^2,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle-\left\langle\psi,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)S_j^2(Y_i,X_i,A_i)\right\} \\ & -E_{(X_{0:N},Y_{1:N})\sim P_{A'_{1:N}}}\left\{\left(\left\langle\psi^2,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle-\left\langle\psi,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)S_j^2(Y_i,X_i,A_i)\right\}\Big|,\end{aligned}$$

it is sufficient to establish Lipschitz continuity, for each i,j, of the function $A_{1:N}\to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle\psi^2,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle-\left\langle\psi,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)S_j^2(Y_i,X_i,A_i)\right\}$. We will show that $S\times h_1$ is Lipschitz continuous provided

$$\sup_{A_{1:N}\in\text{range}(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^4\right\} \quad < \quad \infty, \tag{48}$$

$$\sup_{A_{1:N}\in\text{range}(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{|Y_i|^2\right\} \quad < \quad \infty, \tag{49}$$

$$\sup_{A_{1:N}\in\text{range}(F)} \max_l\left|\frac{1}{X_i^{(l)}(3)-A_i(2)}\right| \quad < \quad \infty, \tag{50}$$

for $1\le i\le N$, which is implied by (42), (43), together with the observation model (31).

Lipschitz continuity of $\overline{S^2}(A_{1:N})$ We commence by showing that the mapping $A_{1:N}\to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\}$ has bounded partial derivatives and hence is Lipschitz. Consider first $\nabla_{A_j(1)}E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^2(Y_i,X_i,A_i)\right\}$. For $j\neq i$,

$$\begin{aligned}&\frac{\partial E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^2(Y_i,X_i,A_i)\right\}}{\partial A_j(1)} \\ =\quad & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^2(Y_i,X_i,A_i)S_1(Y_j,X_j,A_j)\right\} \\ =\quad & 0.\end{aligned}$$

For $j>i$, the result is follows by conditioning on $X_{0:j},Y_{1:j-1}$. For $j<i$, the result follows by conditioning on $X_{0:i},Y_{1:j-1},Y_{j+1:i}$.

For $j = i$,

$$\frac{\partial E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^2(Y_i,X_i,A_i)\right\}}{\partial A_i(1)} = E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^3(Y_i,X_i,A_i)\right\} + E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\frac{\partial S_1^2(Y_i,X_i,A_i)}{\partial A_i(1)}\right\}.$$

Note that $E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_1^3(Y_i,X_i,A_i)\right\} = 0$ as odd central moments of a Gaussian are zero.

$$\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\frac{\partial S_1^2(Y_i,X_i,A_i)}{\partial A_i(1)}\right\} \\
= \ & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{2S_1(Y_i,X_i,A_i)\frac{\partial S_1(Y_i,X_i,A_i)}{\partial A_i(1)}\right\} \\
= \ & 4E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\frac{-\sigma_Y^{-2}}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^3}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\frac{1}{(X_i(3)-A_i(2))^3}\right\} \\
\leq \ & 4\sigma_Y^{-2}E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^3\right\}.
\end{aligned}$$

Similar arguments apply to show all partial derivatives of $E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{S_j^2(Y_i,X_i,A_i)\right\}$ are bounded provided (47) holds.

Lipschitz continuity of $\overline{S\times h_1}(A_{1:N})$ It is sufficient to establish Lipschitz continuity, for each i, j, of the function
$A_{1:N} \to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle\psi^2,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle\psi,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)S_j^2(Y_i,X_i,A_i)\right\}$, with domain range$(F)$. In doing so, we require first the following bounds.

For $j = 1, 2$,

$$\left|S_j(y_i,X_i^{(l)},a_i)\right| \leq \sigma_Y^{-2}\left|\frac{1}{X_i^{(l)}(3)-a_i(2)}\right|(|y_i|+D),$$

where constant D is the bound on the function atan. Thus,

$$\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\max_i\left|S_1(Y_l,X_l^{(i)},A_l)\right|\right)^2\right\} \\
\leq \ & \sigma_Y^{-4}\max_i\left|\frac{1}{X_l^{(i)}(3)-A_l(2)}\right|^2 E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{(|Y_l|+D)^2\right\}.
\end{aligned}$$

Also,

$$\begin{aligned}
& \frac{\partial}{\partial a_i(j)}\left\langle\psi,\hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle, \\
= \ & \left\langle\psi S_j(y_i,\cdot,a_i),\hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle - \left\langle\psi,\hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle\left\langle S_j(y_i,\cdot,a_i),\hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle \\
\leq \ & 2\max_l\left|\psi(X_{0:N}^{(l)})\right| \times \max_l\left|S_j(y_i,X_i^{(l)},a_i)\right|
\end{aligned}$$

for $j=1,2$, $i=1,\ldots,N$. Thus,

$$
\begin{aligned}
& \frac{\partial}{\partial a_i(j)}\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle^2\right) \\
= \; & \frac{\partial}{\partial a_i(j)}\left\langle \psi^2, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle - 2\left\langle \psi, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle \frac{\partial}{\partial a_i(j)}\left\langle \psi, \hat{\pi}_{0:N}^{(y_{1:N},a_{1:N})}\right\rangle \\
\leq \; & 2\max_l \left|\psi^2(X_{0:N}^{(l)})\right| \times \max_l \left|S_j(y_i, X_i^{(l)}, a_i)\right| \\
& +4\max_l \left|\psi(X_{0:N}^{(l)})\right|^2 \times \max_l \left|S_j(y_i, X_i^{(l)}, a_i)\right| \\
= \; & C\times \max_l \left|S_j(y_i, X_i^{(l)}, a_i)\right|,
\end{aligned}
$$

where the constant C is independent of $(y_{1:N}, a_{1:N})$.

Taking the partial derivative, we have

$$
\begin{aligned}
& \frac{\partial}{\partial A_m(n)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right) S_j^2(Y_i, X_i, A_i)\right\} \\
= \; & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right) S_j^2(Y_i, X_i, A_i) S_n(Y_m, X_m, A_m)\right\} \\
& +2E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right) S_j(Y_i, X_i, A_i)\frac{\partial S_j(Y_i, X_i, A_i)}{\partial A_m(n)}\right\} \\
& +E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\frac{\partial\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)}{\partial A_m(n)} S_j^2(Y_i, X_i, A_i)\right\}.
\end{aligned}
$$

Boundedness of the first term follows from the boundedness of $\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)$ and the boundedness of $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{|S_j(Y_i,X_i,A_i)|^p\right\}$, $p>0$. The second term is bounded because $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|S_j(Y_i,X_i,A_i)\frac{\partial S_j(Y_i,X_i,A_i)}{\partial A_m(n)}\right|\right\} < \infty$. (It can be shown that $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^4\right\} < \infty$ for all i implies $\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{\partial S_j(Y_i,X_i,A_i)}{\partial A_m(n)}\right|^2\right\} < \infty$ for all i,j,m,n.) To bound the final term, we use

$$
\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{\partial\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)}{\partial A_m(n)} S_j^2(Y_i, X_i, A_i)\right|\right\} \\
\leq \; & E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{\partial\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)}{\partial A_m(n)}\right|^2\right\}^{\frac{1}{2}} \\
& \times E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|S_j^2(Y_i, X_i, A_i)\right|^2\right\}^{\frac{1}{2}},
\end{aligned}
$$

where

$$\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{\partial\left(\left\langle\psi^2,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle-\left\langle\psi,\hat{\pi}_{0:N}^{(Y_{1:N})}\right\rangle^2\right)}{\partial A_m(n)}\right|^2\right\} \\
\leq\ & C\times E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left(\max_l\left|S_n(Y_m,X_m^{(l)},A_m)\right|\right)^2\right\} \\
\leq\ & C\times\sigma_Y^{-4}\max_l\left|\frac{1}{X_m^{(l)}(3)-A_m(2)}\right|^2\times E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{(|Y_m|+D)^2\right\}.
\end{aligned}$$

So, we need (49) and (50).

4.1.3 Assumption 3

We need to verify Assumption 3 with

$$H_{k+1}=-\Gamma(b_k)\nabla F(\theta_k)\left(h_{1,F(\theta_k)}(\omega_{k+1})+h_{2,F(\theta_k)}(\omega_{k+1})-S_{F(\theta_k)}(\omega_{k+1})b_k\right).$$

We will show Assumption 3 is satisfied if for all $1\leq i\leq N$, $p>0$,

$$\sup_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left\{\left|\frac{1}{X_i(3)-A_i(2)}\right|^p\right\} < \infty, \tag{51}$$

$$\sup_{A_{1:N}\in\text{range(F)}}\max_l\left|\frac{1}{X_i^{(l)}(3)-A_i(2)}\right| < \infty, \tag{52}$$

$$\sup_{A_{1:N}\in\text{range(F)}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left(|Y_i|^p\right) < \infty. \tag{53}$$

In fact, (53) is satisfied for the bearings only observation model (31). Note that $|Y_i|^p\leq 1+|Y_i|^{\lceil p\rceil\times 2}$ where $\lceil p\rceil$ denotes the smallest integer greater than or equal to p. Thus, it is sufficient to verify (53) for positive even integers. We will need the following inequality (see A.N. Shiryaev, Probability, 2nd Ed., pg.194): if $a,b>0$ and $p\geq 1$, then $(a+b)^p\leq 2^{p-1}(a^p+b^p)$.

$$\begin{aligned}
& E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left(|Y_i|^p\right) \\
\leq\ & 2^{p-1}E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left(\left|Y_i-\operatorname{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\right|^p\right) \\
& +2^{p-1}E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}}\left(\left|\operatorname{atan}\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)\right|^p\right) \\
\leq\ & 2^{p-1}\left(C_p+D\right),
\end{aligned}$$

where C_p follows from the central moment of a Gaussian and D is the bound on the function atan.

We may write

$$\begin{aligned}
|H_{k+1}| \leq\ & |\Gamma(b_k)|\,|\nabla F(\theta_k)|\,\left|h_{1,F(\theta_k)}(\omega_{k+1})\right| \\
& +|\Gamma(b_k)|\,|\nabla F(\theta_k)|\,\left|h_{2,F(\theta_k)}(\omega_{k+1})\right|+|\nabla F(\theta_k)|\,|\Gamma(b_k)b_k|\,\left|S_{F(\theta_k)}(\omega_{k+1})\right| \\
\leq\ & C_1\left|h_{1,F(\theta_k)}(\omega_{k+1})\right|+C_1\left|h_{2,F(\theta_k)}(\omega_{k+1})\right|+C_2\left|S_{F(\theta_k)}(\omega_{k+1})\right|,
\end{aligned}$$

which implies

$$|H_{k+1}|^p\leq 2^{2p-2}C_1^p\left(\left|h_{1,F(\theta_k)}(\omega_{k+1})\right|^p+\left|h_{2,F(\theta_k)}(\omega_{k+1})\right|^p\right)+2^{p-1}C_2^p\left|S_{F(\theta_k)}(\omega_{k+1})\right|^p.$$

It is sufficient to verify
$\sup_k E(|h_{1,F(\theta_k)}(\omega_{k+1})|^p) < \infty$, $\sup_k E(|h_{2,F(\theta_k)}(\omega_{k+1})|^p) < \infty$ and $\sup_k E(|S_{F(\theta_k)}(\omega_{k+1})|^p) < \infty$ for all $p > 0$.

Recall the definition of $h_{1,A_{1:N,k}}(X_{0:N,k+1}, Y_{1:N,k+1})$,

$$\begin{aligned} & h_{1,A_{1:N,k}}(X_{0:N,k+1}, Y_{1:N,k+1})) \\ = & \left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle^2 \right) S_{A_{1:N,k}}(X_{0:N,k+1}, Y_{1:N,k+1})\mathbf{1}. \end{aligned}$$

Thus we have $|h_{1,F(\theta_k)}(\omega_{k+1})| \le C \times |S_{F(\theta_k)}(\omega_{k+1})|$, and is sufficient to verify $\sup_k E(|S_{F(\theta_k)}(\omega_{k+1})|^p) < \infty$.

To establish $\sup_k E(|h_{2,F(\theta_k)}(\omega_{k+1})|^p) < \infty$ for all $p > 0$, is sufficient to show $\sup_k E(|[h_{2,F(\theta_k)}(\omega_{k+1})]_i|^p) < \infty$ for all $p > 0$, i, where $[h_{2,F(\theta_k)}(\omega_{k+1})]_i$ is the i-th component of the vector $h_{2,F(\theta_k)}(\omega_{k+1})$. From the definition of $h_{2,F(\theta_k)}(\omega_{k+1})$, this implies showing

$$\begin{gathered} \sup_k E \left(\left| \left\langle \psi^2 S_j(Y_{i,k+1}, \cdot, A_{i,k}), \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \right.\right. \\ - \left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \left\langle S_j(Y_{i,k+1}, \cdot, A_{i,k}), \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \\ -2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \left\langle \psi S_j(Y_{i,k+1}, \cdot, A_{i,k}), \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \\ \left.\left. +2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle^2 \left\langle S_j(Y_{i,k+1}, \cdot, A_{i,k}), \hat{\pi}_{0:N}^{(Y_{1:N,k+1})} \right\rangle \right|^p \right) \\ \le C \times \sup_k E \left(\max_l \left| S_j(Y_{i,k+1}, X_i^{(l)}, A_{i,k}) \right|^p \right) < \infty, \end{gathered}$$

for $j = 1, 2, i = 1, \ldots, N$, where the constant C depends only on the function ψ and the state samples $\{X_{0:N}^{(l)}\}_l$ used in $\hat{\pi}_{0:N}$. It was established above that for $j = 1, 2$,

$$\max_l \left| S_j(y_i, X_i^{(l)}, a_i) \right| \le \sigma_Y^{-2} \max_l \left| \frac{1}{X_i^{(l)}(3) - a_i(2)} \right| (|y_i| + D),$$

where the constant D is the bound on the atan function. Thus

$$\begin{aligned} & \sup_k E \left(\max_l \left| S_j(Y_{i,k+1}, X_i^{(l)}, A_{i,k}) \right|^p \right) \\ \le & \sigma_Y^{-2p} 2^{p-1} \sup_k E \left(\max_l \left| \frac{1}{X_i^{(l)}(3) - A_{i,k}(2)} \right|^p (|Y_{i,k+1}|^p + D^p) \right) \\ \le & \sigma_Y^{-2p} 2^{p-1} \sup_{A_i \in \text{range(F)}} \max_l \left| \frac{1}{X_i^{(l)}(3) - A_i(2)} \right|^p \sup_k E(|Y_{i,k+1}|^p + D^p) \\ \le & \sigma_Y^{-2p} 2^{p-1} \sup_{A_i \in \text{range(F)}} \max_l \left| \frac{1}{X_i^{(l)}(3) - A_i(2)} \right|^p \sup_{A_{1:N} \in \text{range(F)}} E_{Y_i \sim P_{A_{1:N}}} (|Y_i|^p + D^p). \end{aligned}$$

Thus, we require

$$\begin{aligned} \sup_{A_i \in \text{range(F)}} \max_l \left| \frac{1}{X_i^{(l)}(3) - A_i(2)} \right| & < \infty, \\ \sup_{A_{1:N} \in \text{range(F)}} E_{Y_i \sim P_{A_{1:N}}} (|Y_i|^p) & < \infty, \end{aligned}$$

for $\sup_k E\left(\max_l \left|S_1(Y_{i,k+1}, X_i^{(l)}, A_{i,k})\right|^p\right) < \infty$.

Since $E(|S_{F(\theta_k)}(\omega_{k+1})|^p) = E\left[E_{\omega_{k+1} \sim P_{F(\theta_k)}}(|S_{F(\theta_k)}(\omega_{k+1})|^p)\right]$, Assumption 2c implies $\sup_k E(|S_{F(\theta_k)}(\omega_{k+1})|^p) < \infty$ is also satisfied.

4.1.4 Assumption 4

We need to establish that exists some constant $a > 0$ such that for all $A_{1:N}, b \in R^{2N}$, $b^T\overline{S^2}(A_{1:N})b \geq a|b|^2$. This would follow if

$$\inf_{j,i} \inf_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{S_j^2(Y_i, X_i, A_i)\right\} > 0.$$

So, we require

$$\begin{aligned} &\inf_{i} \inf_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{S_1^2(Y_i, X_i, A_i)\right\} \\ = \; &\inf_{i} \inf_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \sigma_Y^{-2} \frac{\frac{1}{(X_i(3)-A_i(2))^2}}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2} \right\} > 0, \end{aligned}$$

$$\begin{aligned} &\inf_{i} \inf_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{S_2^2(Y_i, X_i, A_i)\right\} \\ = \; &\inf_{i} \inf_{A_{1:N}} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \sigma_Y^{-2} \frac{\left[\frac{X_i(1)-A_i(1)}{(X_i(3)-A_i(2))^2}\right]^2}{\left[1+\left(\frac{X_i(1)-A_i(1)}{X_i(3)-A_i(2)}\right)^2\right]^2} \right\} > 0. \end{aligned}$$

4.2 Verifying conditions for the convergence of the slow timescale

We will show that conditions (a) and (b) of Corollary 10 are satisfied. The assumptions necessary for the convergence of the fast timescale imply condition (c) is satisfied (see Corollary 9).

4.2.1 Corollary 10, condition (a)

By Remark 11, assuming F has bounded first and second order derivatives, the Lipschitz condition on $\nabla(J\circ F)$ is satisfied provided $|\bar{h}(F(\theta))|$ is bounded and $\nabla J = \bar{h}$ is Lipschitz when its domain is restricted to range(F).

To show $|\bar{h}(F(\theta))|$ is bounded, it is sufficient to show each component of the vector valued functions $\bar{h}_1(F(\theta))$ and $\bar{h}_2(F(\theta))$ is bounded. Likewise, to show $\bar{h} = \bar{h}_1 + \bar{h}_2$ is Lipschitz, it is sufficient to show each component of the vector valued functions $\bar{h}_1$ and $\bar{h}_2$ is Lipschitz.

For the trajectory planning problem, the components of $\bar{h}_1(A_{1:N})$ are the functions

$$A_{1:N} \rightarrow E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle^2 \right) S_j(Y_i, X_i, A_i) \right\}$$

where $j = 1, 2,\ i = 1, \ldots, N$. Since $\left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle^2 \right)$ is bounded, this function is bounded provided $\sup_{A_{1:N}\in \text{range}(F)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \{|S_j(Y_i, X_i, A_i)|\} < \infty$, which has been verified in Assumption 2c. As for the Lipschitz continuity, we may bound

$$\frac{\partial}{\partial A_m(n)} E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \left\{ \left(\left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle - \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle^2 \right) S_j(Y_i, X_i, A_i) \right\}$$

using similar arguments that were used to establish the Lipschitz continuity of $\overline{S \times h_1}(A_{1:N})$.

The components of $\bar{h}_2(A_{1:N})$ are the functions

$$
\begin{aligned}
A_{1:N} \to E_{(X_{0:N},Y_{1:N})\sim P_{A_{1:N}}} \Big\{ & \left\langle \psi^2 S_j(Y_i,\cdot,A_i), \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \\
& - \left\langle \psi^2, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \left\langle S_j(Y_i,\cdot,A_i), \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \\
& -2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \left\langle \psi S_j(Y_i,\cdot,A_i), \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \\
& +2 \left\langle \psi, \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle^2 \left\langle S_j(Y_i,\cdot,A_i), \hat{\pi}_{0:N}^{(Y_{1:N})} \right\rangle \Big\},
\end{aligned}
$$

where $j = 1, 2$, $i = 1, \ldots, N$. Showing boundedness follows similar arguments used to establish Assumption 3. As for the Lipschitz continuity, we may bound its partial derivatives using similar arguments used in verifying Assumption 2b (and 3).

4.2.2 Corollary 10, condition (b)

In verifying Assumption 3, it was shown that $\sup_k E\left(|H_k|^2\right) < \infty$.

In Remark 12, it was stated that one may verify

$$
\begin{aligned}
\sup_{\tilde{\theta}\in\mathrm{range}(F)} E_{\omega\sim P_{\tilde{\theta}}} \left\{ \left| h_{1,\tilde{\theta}}(\omega) - \bar{h}_1(\tilde{\theta}) \right|^2 \right\} &< \infty, \\
\sup_{\tilde{\theta}\in\mathrm{range}(F)} E_{\omega\sim P_{\tilde{\theta}}} \left\{ \left| h_{2,\tilde{\theta}}(\omega) - \bar{h}_2(\tilde{\theta}) \right|^2 \right\} &< \infty, \\
\sup_{\tilde{\theta}\in\mathrm{range}(F)} E_{\omega\sim P_{\tilde{\theta}}} \left\{ \left| S_{\tilde{\theta}}(\omega) \right|^2 \right\} &< \infty,
\end{aligned}
$$

in lieu of (17)-(19). To verify condition (a) of Corollary 10, it was shown that $\sup_{\tilde{\theta}\in\mathrm{range}(F)} \bar{h}_i(\tilde{\theta}) < \infty$, $i = 1, 2$. In verifying Assumption 3, it was shown that $\sup_{\tilde{\theta}\in\mathrm{range}(F)} E_{\omega\sim P_{\tilde{\theta}}} \left\{ \left| h_{i,\tilde{\theta}}(\omega) \right|^2 \right\} < \infty$, $i = 1, 2$, while in Assumption 2c, it was shown that $\sup_{\tilde{\theta}\in\mathrm{range}(F)} E_{\omega\sim P_{\tilde{\theta}}} \left\{ \left| S_{\tilde{\theta}}(\omega) \right|^2 \right\} < \infty$.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, "Gradient Convergence in Gradient Methods with Errors," *SIAM Journal on Optimization*, Vol. 10, No. 3, pp. 627-642, 2000.

[2] A. Doucet, N.J. Gordon and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Processing*, vol. 49, no. 3, pp. 613–624, Mar. 2001.

[3] A. Doucet, J.F.G. de Freitas and N.J. Gordon, *Sequential Monte Carlo methods in practice*. New York: Springer, 2001.

[4] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Applications*. Academic Press, 1980.

[5] V. R. Konda and J. N. Tsitsiklis, "Linear Stochastic Approximation Driven by Slowly Varying Markov Chains," *Systems and Control Letters*, vol. 50, No. 2, pp. 95-102, 2003.

[6] B. T. Polyak, "Pseudogradient adaptation and training algorithms," *Automation and Remote Control*, vol. 34, No. 3, pp. 377-397, 1973.

[7] S. S. Singh, N. Kantas, B. Vo, A. Doucet and R. Evans, "Simulation-Based Optimal Sensor Scheduling with Application to Observer Trajectory Planning," CUED/F-INFENG/TR 521, University of Cambridge, Department of Engineering, March 2005.