2018 IEEE 14th International Conference on Control and Automation (ICCA)
June 12-15, 2018. Anchorage, Alaska, USA

Estimation of Random Mobility Models using the Expectation-Maximization Method

Tao Li[1], Yan Wan[1], Mushang Liu[1], and Frank L. Lewis[2]

Abstract— **Random mobility models (RMMs) capture the statistical movement characteristics of mobile agents, and have been widely used for the evaluation and design of mobile wireless networks. In many RMMs, the movement characteristics are captured as stochastic processes constructed using two types of independent random variables. The first type describes the movement characteristics for each maneuver, and the second type describes how often the maneuvers are switched. In this paper, we develop a generic method to estimate RMMs that are composed of these two types of random variables. In particular, we formulate the dynamics of movement characteristics generated by the two types of random variables as a special Jump Markov system, and develop an estimation method based on the Expectation-Maximization principle.**

I. INTRODUCTION

Random mobility models (RMMs) capture the movement characteristics of mobile agents, and have been widely used to evaluate the performance of mobile wireless networks [1]. Many RMMs have been developed, ranging from the basic ones (e.g., Random Direction and Random Walk) to more sophisticated ones designed for specific vehicle types (e.g., ground vehicles and airborne vehicles [2]) and vehicle movement patterns with specific constraints (e.g., [3]). Additional examples of RMMs can be found in several survey papers [1], [4].

Despite the wide variability, many RMMs capture the movement characteristics as stochastic processes constructed using two types of independent random variables. Type 1 describes the movement characteristics for each maneuver. Type 2 describes how often the maneuvers are switched. For example, paper [2] developed the smooth-turn mobility (ST) RMM to capture the smooth movement of fixed-wing unmanned aerial vehicles (UAVs). The ST RMM is composed of a sequence of switching turning maneuvers, in which the turning radius in each maneuver is captured as a type 1 random variable, and the duration of each maneuver is captured as a type 2 random variable.

In order for the RMMs to generate realistic movement characteristics, parameters in the two types of random variables need to be properly specified using real trajectory data. Most of the existing estimation methods (see, e.g., [5], [6]) were developed only for specific RMMs and lack the flexibility to be applied to other RMMs. In most of these studies, the movement characteristics are captured by only one random variable and hence the parameters for the random variable can be simply estimated through a direct fitting of the associated distribution with the observed trajectory data statistics. However, for the RMMs that we consider, estimating the parameters in one type by simply fitting the distribution with the trajectory data would result in inaccurate estimates because the trajectory is determined by the two types of random variables jointly.

To the best of our knowledge, only a few studies have been devoted to estimating parameters in the two types of random variables using trajectory data. Paper [7] developed a heuristic method to estimate both types of random variables in a 2-Dimensional (2D) ST RMM. Specifically, the turning radius and switching points are first estimated from trajectory data by heuristically balancing among multiple criteria (e.g., degree of correlation and estimation error statistics). Then, the parameters in each type of random variables are estimated using the corresponding statistics. Paper [8] adopted a similar approach to estimate the random variable that determines the travel pause time of cell phone users. One drawback of this approach is that parameters are estimated based on one possible (heuristic) separation of the trajectory into maneuver sessions out of a large number of possibilities. Large errors can arise when the noise level in the trajectory data is high.

This study is focused on estimating parameters in the two types of random variables using trajectory data. In contrast to [7], [8] which estimate the parameters in two types of random variables separately for specific movement characteristics of specific RMMs, the originality of our work is two-fold. First, we estimate parameters simultaneously by considering all possible separations of trajectory data into maneuver sessions. In other words, our method avoids the drawback of the methods in [7], [8], and provides more reliable and accurate estimates. Second, our estimation method is general in that it is not restricted to a specific movement characteristics of a specific RMM.

We show that the parameter estimation problem falls into the category of estimating a Jump Markov Linear System (JMLS). The latter is NP-hard. We adopt the Expectation-Maximization (EM) algorithm to estimate the parameters. The algorithm is an iterative procedure, which generates maximum likelihood parameter estimates. It has been widely used in estimating JMLS (see, for example, [9], [10]). However, the existing work on the estimation of the JMLS (e.g., [11], [12]) focuses on the direct estimation of states, model coefficients, or transition matrix of the modulating Markov chain. In our study, the model coefficients and transition matrix are functions of the parameters that describe the two

*We thank the National Science Foundation under grant numbers 1714519 and 1730675 for the support of this work.

[1]Tao Li, Yan Wan, Mushang Liu are with the Department of Electrical Engineering, University of Texas at Arlington. Email: tao.li2@uta.edu, yan.wan@uta.edu (contact author), mushuang.liu@mavs.uta.edu.

[2] Frank L. Lewis is with UTA Research Institute, University of Texas at Arlington, Fort Worth, Texas, USA., and Qian Ren Consulting Professor, State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China. Email: lewis@uta.edu.

types of random variables. We apply the EM algorithm to estimate these parameters.

II. Problem Formulation

In this section, we formulate the parameter estimation problem as a special JMLS estimation problem. We consider a general modeling framework:

$$\begin{aligned} x_t &= A(z_t, r_t)x_{t\text{-}1} + F(z_t, r_t)u_t + H(z_t, r_t), \\ y_t &= C(z_t, r_t)x_t + D(z_t, r_t)w_t + G(z_t, r_t)u_t, \end{aligned} \tag{1}$$

where x_t is the system state, y_t is the measurement of x_t, and $r_t \in \{s_0, s_1\}$ is a discrete Markov chain of two states.

$$Z_t = \begin{cases} Z_{t\text{-}1}, & \text{if } r_t = s_0, \\ \hat{Z}_t, & \text{if } r_t = s_1. \end{cases} \tag{2}$$

where $\hat{Z}_t$ is a random variable with the pdf $f(\hat{Z}_t = \hat{z}_t)$ denoted as $f(\hat{z}_t; \theta_1)$ with parameter θ_1. The transition matrix is given as follows:

$$p_{s_1,s_1}(t) = P(r_t = s_1 | r_{t\text{-}1} = s_1) = g_1(\theta_2), \tag{3}$$

$$p_{s_0,s_1}(t) = P(r_t = s_1 | r_{t\text{-}1} = s_0) = g_0(\theta_2), \tag{4}$$

where θ_2 is a parameter, and g_0 and g_1 are functions of θ_2. θ_1 and θ_2 both can be vectors. w_t is zero-mean white Gaussian noise with known variance Σ_w and we further assume that $D(z_t, r_t)\Sigma_w D(z_t, r_t)' > 0$. u_t is a known deterministic input. The matrices A, H, C, D, F, G are called model coefficients in this study and are all known functions of z_t and r_t. Furthermore, we assume that $A(\cdot, s_1)$ is a zero matrix to guarantee that the movement characteristics before and after a switch are independent.

The focus of this study is to estimate the parameters $\theta = [\theta_1, \theta_2]'$. θ_1 determines the type 1 random variable for the movement characteristics of each maneuver, and θ_2 determines the type 2 random variable for the maneuver switching behavior.

Estimating parameters of the system described by (1)-(4) is related to estimating a JMLS. In the literature, a significant amount of work has been focused on estimating the latter. The mostly used JMLS is of the following form:

$$\begin{aligned} x_t &= A(\overline{r}_t)x_{t\text{-}1} + B(\overline{r}_t)v_t + F(\overline{r}_t)u_t, \\ y_t &= C(\overline{r}_t)x_t + D(\overline{r}_t)w_t + G(\overline{r}_t)u_t, \end{aligned} \tag{5}$$

where v_t is a zero-mean white Gaussian noise, $\overline{r}_t \in \{1, 2, \ldots U\}$ is the state of a discrete Markov chain with the transition matrix $P_{\overline{r}}$. One important difference between the two systems is as follows. The model coefficients and transition matrix (both of which are usually called model parameters in the literature) in (5) are functions of $\overline{r}_t$ and their values are typically in finite sets. In contrast, the model coefficients and transition matrix in (1) are functions of r_t and z_t, which correspond to the two types of random variables. Another difference is that we introduce a term $H(z_t, r_t)$ in (1) to create a switch to the movement characteristic that is independent from previous ones (i.e., $x_{1:t\text{-}1}$).

Estimating a JMLS can be NP-hard as the number of possible realizations of $\overline{r}_t$ grows exponentially with the size of states. While most of the existing estimation studies focus on estimating the states (i.e., $x_t, \overline{r}_t$), model coefficients (i.e., matrices $A(\overline{r}_t), B(\overline{r}_t), C(\overline{r}_t), D(\overline{r}_t), F(\overline{r}_t), G(\overline{r}_t)$) and transition matrix (i.e., $P_{\overline{r}_t}$), this study focuses on estimating the parameter θ, which determines the states, model coefficients, and transition matrix in (1).

III. Estimation methodology

In this section, we present the methodology to generate a maximum likelihood estimate of the parameter θ.

Let $\overline{Y} = [Y^1, Y^2, \ldots, Y^N]$ be the set of N mutually independent measurement experiments. For each $s \in \{1, ..., N\}$, $Y^s = [y_0^s, y_1^s \ldots, y_{L_s}^s, y_{L_s}^s]'$ is composed of $L_s + 1$ measurements of y_t, where $t = 0, ..., L_s$. $\overline{R} = [R^1, R^2, \ldots, R^N]$ is the set of actual Markov chain states corresponding to $\overline{Y}$, where $R^s = [r_0^s, r_1^s \ldots, r_{L_s}^s, r_{L_s}^s]'$. In this study, we assume that $r_0^s = s_1$, $s = 1, ..., N$.

The maximum likelihood estimate of θ is given by

$$\arg_\theta \max P(\theta | \overline{Y}). \tag{6}$$

where

$$P(\theta | \overline{Y}) = \frac{P(\overline{Y}, \theta)}{P(\overline{Y})} = \frac{P(\overline{Y} | \theta) P(\theta)}{P(\overline{Y})}. \tag{7}$$

Directly calculating $P(\overline{Y}|\theta)$ is difficult as the computations of high-dimensional integrals are not always tractable analytically. In the following, we apply the EM to estimate θ. The main idea of the EM is to treat $\overline{Y}$ as incomplete data and introduce a latent variable R (i.e., the Markov chain states) for which the joint likelihood $P(\overline{Y}, R|\theta)$ is available and easier to evaluate. The EM solves for θ that maximizes the expected log-likelihood of the complete data. More specifically, the algorithm estimates θ using two iterative steps: E-step and M-step. It first calculates the expected value of the log-likelihood of the latent variables for a given parameter estimate (E-step) and then updates the parameter estimate by maximizing the expected value from E-step (M-step). The process is repeated until the convergence is reached.

EM formulation: By introducing the latent variable R representing the Markov chain states, we solve the following problem for θ.

$$\arg_\theta \max P(\overline{Y}, R, \theta). \tag{8}$$

Let

$$\begin{aligned} L(\overline{Y}, R, \theta) &= \log(P(\theta, R, \bar{Y})) \\ &= \log(P(\theta)) + \log(P(\overline{Y}, R|\theta)), \end{aligned} \tag{9}$$

where $P(\theta)$ is the prior distribution of θ and assumed to be uniform in this study.

The EM algorithm is summarized as follows:

E-Step:

$$\phi(\theta, \theta^l) = E_R(L(\overline{Y}, R, \theta) | \overline{Y}, \theta^l), \tag{10}$$

where θ^l is the parameter estimated from previous step l.
M-Step:

$$\theta^{l+1}=\arg_{\theta}\max(\phi(\theta,\theta^l)) \quad s.t. \;\; \theta_L \leq \theta \leq \theta_U, \tag{11}$$

The two steps are repeated until the convergence is achieved.

Different from the method in [7], [8], the EM method estimates the parameter θ by considering all the possible realizations of R for $\overline{Y}$ (via the expected value).

A. E-step

In the E-step, we calculate the function $\phi(\theta,\theta^l)$. Since $\log(P(\theta))$ is a constant with respect to R, we have

$$\begin{aligned}\phi(\theta,\theta^l)=&E_R(L(\overline{Y},R,\theta)|\overline{Y},\theta^l)\\ =&\log(P(\theta))+E_R(\log(P(\overline{Y},R|\theta)|\overline{Y},\theta^l))\\ =&\log(P(\theta))+\Sigma_s\Sigma_i p_i^s\log(P(Y^s,R_i^s|\theta)),\end{aligned} \tag{12}$$

where p_i^s=$P(R_i^s|Y^s,\theta^l)$, R_i^s is the ith possible realization of the discrete Markov chain $\{r_t\}_{t=0}^{L_s}$, and $(R_i^s)_t$ represents the value of the Markov chain at t in R_i^s. The second term in (12) can be rewritten as

$$\begin{aligned}\Sigma_s\Sigma_i p_i^s\log(P(Y^s,R_i^s|\theta))=&\Sigma_s\Sigma_i p_i^s\log(P(Y^s|R_i^s,\theta))\\ &+\Sigma_s\Sigma_i p_i^s\log(P(R_i^s|\theta)).\end{aligned} \tag{13}$$

For the first term in (13), we have

$$\begin{aligned}&\Sigma_s\Sigma_i p_i^s\log(P(Y^s|R_i^s,\theta))\\ &=\Sigma_s\Sigma_{t=0}^{L_s}\Sigma_i p_i^s\log(P(y_t^s|y_{0:t\text{-}1}^s,R_i^s,\theta))\\ &=\Sigma_s\Sigma_{t=0}^{L_s}[\Sigma_{(i,r_t^s=s_1)}p_i^s\log(P(y_t^s|r_t^s=s_1,\theta))\\ &\;+\Sigma_{n=0}^{t\text{-}1}\Sigma_{(i,r_{t\text{-}(n+1):t}^s=S_{n+2})}p_i^s\\ &\;\times\log(P(y_t^s|y_{0:t\text{-}1}^s,r_{t\text{-}(n+1):t}^s=S_{n+2},\theta))]\\ &=\Sigma_{s,t}[\log(P(y_t^s|r_t^s=s_1,\theta))(\Sigma_{(i,r_t^s=s_1)}p_i^s)\\ &\;+\Sigma_{n=0}^{t\text{-}1}\log(P(y_t^s|y_{0:t\text{-}1}^s,r_{t\text{-}(n+1):t}^s=S_{n+2},\theta))\\ &\;\times(\Sigma_{(i,r_{t\text{-}(n+1):t}^s=S_{n+2})}p_i^s)]\\ &=\Sigma_{s,t}[\log(P(y_t^s|r_t^s=s_1,\theta))P(r_t^s=s_1|Y^s,\theta)\\ &\;+\Sigma_{n=0}^{t\text{-}1}\log(P(y_t^s|y_{0:t\text{-}1}^s,r_{t\text{-}(n+1):t}^s=S_{n+2},\theta))\\ &\;\times P(r_{t\text{-}(n+1):t}^s=S_{n+2}|Y^s,\theta^l)],\end{aligned} \tag{14}$$

where

$$\begin{aligned}r_{t\text{-}(n+1):t}^s=&[r_{t\text{-}(n+1)},r_{t\text{-}n},\ldots,r_t]_{1\times(n+2)},\\ S_{n+2}=&[s_1,s_0,\ldots,s_0]_{1\times(n+2)}.\end{aligned}$$

The second term in (13) can be expressed as

$$\begin{aligned}&\Sigma_s\Sigma_i p_i^s\log(P(R_i^s|\theta))\\ &=\Sigma_s\Sigma_i p_i^s\log[P(r_0^s=s_1|\theta)\Pi_{t=1}^{L_s}P(r_t^s|r_{t\text{-}1}^s,\theta)]\\ &=\Sigma_s\Sigma_{t=1}^{L_s}\Sigma_i p_i^s\log(P(r_t^s|r_{t\text{-}1}^s,\theta))\\ &+\Sigma_s\Sigma_i p_i^s\log(P(r_0^s=s_1|\theta)).\end{aligned} \tag{15}$$

The second term in (15) is zero as the initial r_0^s=s_1 is independent of θ. Then, we have

$$\begin{aligned}&\Sigma_s\Sigma_i p_i^s\log(P(R_i^s|\theta))\\ &=\Sigma_s\Sigma_{t=1}^{L_s}[\Sigma_{r_{t\text{-}1:t}^s}\log(P(r_t^s|r_{t\text{-}1}^s,\theta))\\ &\times(\Sigma_i p_i^s 1_{\{(R_i^s)_{t\text{-}1:t}=r_{t\text{-}1:t}^s\}})]\\ &=\Sigma_s\Sigma_{t=1}^{L_s}\Sigma_{r_{t\text{-}1:t}^s}\log(P(r_t^s|r_{t\text{-}1}^s,\theta))\\ &\times P(r_{t\text{-}1:t}^s|Y^s,\theta^l),\end{aligned} \tag{16}$$

where

$$1_{\{(R_i^s)_{t\text{-}1:t}=r_{t\text{-}1:t}^s\}}=\begin{cases}1, & \text{if } (R_i^s)_{t\text{-}1:t}=r_{t\text{-}1:t}^s,\\ 0, & \text{otherwise.}\end{cases} \tag{17}$$

The probabilities $P(r_t^s|Y^s,\theta^l)$ and $P(r_{t\text{-}(n+1):t}^s=S_{n+2}|Y^s,\theta^l)$ in (14) and $P(r_{t\text{-}1}^s,r_t^s|Y^s,\theta^l)$ in (16) can be calculated using the backward-forward algorithm. All the following probabilities in this section are conditioned on θ^l. or notation simplicity we dropped the notation for the conditional probability on θ^l as long as it does not cause confusion. To facilitate the backward-forward operation, we define

$$a_t^s(r_t^s)=P(y_{0:t}^s,r_t^s), \tag{18}$$

$$b_t^s(r_t^s)=P(y_{t+1:L_s}^s|y_{0:t}^s,r_t^s). \tag{19}$$

We first calculate $a_t^s(r_t^s)$ and $b_t^s(r_t^s)$ before we calculate the probabilities in (14) and (16).

Recursive calculation of a_t^s: $a_t^s,t\geq 1$ is expressed using the following recursive relationship:

$$\begin{aligned}a_t^s=&P(y_t^s|y_{0:t\text{-}1}^s,r_t^s)P(y_{0:t\text{-}1}^s,r_t^s)\\ =&P(y_t^s|y_{0:t\text{-}1}^s,r_t^s)[\Sigma_{r_{t\text{-}1}^s}P(r_t^s|y_{0:t\text{-}1}^s,r_{t\text{-}1}^s)P(y_{0:t\text{-}1}^s,r_{t\text{-}1}^s))]\\ =&P(y_t^s|y_{0:t\text{-}1}^s,r_t^s)[\Sigma_{r_{t\text{-}1}^s}P(r_t^s|r_{t\text{-}1}^s)a_{t\text{-}1}^s],\end{aligned} \tag{20}$$

Based on (3) and (4) and the fact that the state of the Markov chain at t does not depend on previous measurements, in (20), we have $P(r_t^s|y_{0:t\text{-}1}^s,r_{t\text{-}1}^s)$=$P(r_t^s|r_{t\text{-}1}^s)$. $P(r_t^s|r_{t\text{-}1}^s)$ in (20) is given by (3) and (4).

To calculate $P(y_t^s|y_{0:t\text{-}1}^s,r_t^s)$ in (20), we consider the following two cases based on the value of r_t^s:

Case 1: r_t^s=s_1. Since a new $\hat{z}_t$ will be generated by (2) independently from previous states and parameters at time t, it is straightforward that

$$P(y_t^s|y_{0:t\text{-}1}^s,r_t^s=s_1)=P(y_t^s|r_t^s=s_1). \tag{21}$$

The probability is calculated using the following general result with n=-1.

$$\begin{aligned}&P(y_{t\text{-}(n+1):t}^s|r_{t\text{-}(n+1):t}^s=S_{n+2})\\ &=\int P(y_{t\text{-}(n+1):t}^s|\hat{z}_{t\text{-}(n+1)},r_{t\text{-}(n+1):t}^s=S_{n+2})\\ &\times f(\hat{z}_{t\text{-}(n+1)};\theta_1^l)d\hat{z}_{t\text{-}(n+1)}\\ &=\int\Pi_{j=t\text{-}(n+1)}^t P(y_j^s|\hat{z}_{t\text{-}(n+1)},r_{t\text{-}(n+1):t}^s=S_{n+2})\\ &\times f(\hat{z}_{t\text{-}(n+1)};\theta_1^l)d\hat{z}_{t\text{-}(n+1)},\end{aligned} \tag{22}$$

where

$$y_j^s|(\hat{z}_{t\text{-}(n+1)}, r^s_{t\text{-}(n+1):t}{=}S_{n+2}) \sim N(\mu_j^s, D(\hat{z}_{t\text{-}(n+1)}, r_j)\Sigma_w D'(\hat{z}_{t\text{-}(n+1)}, r_j)), \quad (23)$$

and

$$\mu_j^s{=}G(\hat{z}_{t\text{-}(n+1)}, s_0)u_j^s{+}C(\hat{z}_{t\text{-}(n+1)}, s_0) \times [\Sigma_{l=0}^{j\text{-}(t\text{-}n)} A^l(\hat{z}_{t\text{-}(n+1)}, s_0)F(\hat{z}_{t\text{-}(n+1)}, s_0)u^s_{j\text{-}l} {+}A^{j\text{-}(t\text{-}n\text{-}1)}(\hat{z}_{t\text{-}(n+1)}, s_0) \times [H(\hat{z}_{t\text{-}(n+1)}, s_1){+}F(\hat{z}_{t\text{-}(n+1)}, s_1)u^s_{t\text{-}(n+1)}]]. \quad (24)$$

This completes the calculations for Case 1.

Case 2: $r_t^s{=}s_0$.

$$P(y_t^s|y^s_{0:t\text{-}1}, r_t^s{=}s_0) {=}\Sigma_{n=0}^{t\text{-}1} P(y_t^s, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t\text{-}1}, r_t^s{=}s_0) {=}[\Sigma_{n=0}^{t\text{-}1} P(y_t^s|y^s_{0:t\text{-}1}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r_t^s{=}s_0) \times P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t\text{-}1}, r_t^s{=}s_0)]. \quad (25)$$

In the following, we calculate and $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t\text{-}1}, r_t^s{=}s_0)$ in (25).

To calculate the first probability $P(y_t^s|y^s_{0:t\text{-}1}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r_t^s{=}s_0)$, we have

$$P(y_t^s|y^s_{0:t\text{-}1}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r_t^s{=}s_0) {=}P(y_t^s|y^s_{0:t\text{-}1}, r^s_{t\text{-}(n+1):t}{=}S_{n+2}) {=}\frac{P(y^s_{t\text{-}(n+1):t}|r^s_{t\text{-}(n+1):t}{=}S_{n+2})}{P(y^s_{t\text{-}(n+1):t\text{-}1}|r^s_{t\text{-}(n+1):t}{=}S_{n+2})} {=}\frac{P(y^s_{t\text{-}(n+1):t}|r^s_{t\text{-}(n+1):t}{=}S_{n+2})}{P(y^s_{t\text{-}(n+1):t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1})}. \quad (26)$$

The second equation in (26) holds because r_t^s is independent of $y^s_{0:t\text{-}1}$. The probabilities in (26) are calculated using (22).

For the second probability, we have

$$P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t\text{-}1}, r_t^s{=}s_0) {=}\frac{P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t}{=}S_{n+2})}{P(y^s_{0:t\text{-}1}|r_t^s{=}s_0)} \times P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r_t^s{=}s_0). \quad (27)$$

where

$$P(y^s_{0:t\text{-}1}|r_t^s{=}s_0){=}\Sigma_{n=0}^{t\text{-}1} P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t}{=}S_{n+2}) \times P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r_t^s{=}s_0) \quad (28)$$

Now we calculate the probability $P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t}{=}S_{n+2})$. As r_t^s is independent of $y^s_{0:t\text{-}1}$, we have

$$P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t}{=}S_{n+2}) {=}P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}), {=}P(y^s_{0:t\text{-}(n+2)}|r^s_{t\text{-}(n+1)}{=}s_1) \times P(y^s_{t\text{-}(n+1):t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}). \quad (29)$$

For the first term in (29), we have

$$P(y^s_{0:t\text{-}(n+2)}|r^s_{t\text{-}(n+1)}{=}s_1) {=}\Sigma_{r^s_{t\text{-}(n+2)}} P(y^s_{0:t\text{-}(n+2)}|r^s_{t\text{-}(n+2)}, r^s_{t\text{-}(n+1)}{=}s_1) \times P(r^s_{t\text{-}(n+2)}|r^s_{t\text{-}(n+1)}{=}s_1) {=}\Sigma_{r^s_{t\text{-}(n+2)}} P(y^s_{0:t\text{-}(n+2)}|r^s_{t\text{-}(n+2)})P(r^s_{t\text{-}(n+2)}|r^s_{t\text{-}(n+1)}{=}s_1) {=}\Sigma_{r^s_{t\text{-}(n+2)}} P(y^s_{0:t\text{-}(n+2)}, r^s_{t\text{-}(n+2)})\frac{P(r^s_{t\text{-}(n+1)}{=}s_1|r^s_{t\text{-}(n+2)})}{P(r^s_{t\text{-}(n+1)}{=}s_1)} {=}\Sigma_{r^s_{t\text{-}(n+2)}} a^s_{t\text{-}(n+2)}\frac{P(r^s_{t\text{-}(n+1)}{=}s_1|r^s_{t\text{-}(n+2)})}{P(r^s_{t\text{-}(n+1)}{=}s_1)}. \quad (30)$$

Therefore, (27) becomes

$$P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t\text{-}1}, r_t^s{=}s_0) {=}\left(\Sigma_{r^s_{t\text{-}(n+2)}} a^s_{t\text{-}(n+2)}\frac{P(r^s_{t\text{-}(n+1)}{=}s_1|r^s_{t\text{-}n\text{-}2})}{P(r^s_{t\text{-}(n+1)}{=}s_1)}\right) \times P(y^s_{t\text{-}(n+1):t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}) \times\frac{P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r_t^s{=}s_0)}{P(y^s_{0:t\text{-}1}|r_t^s{=}s_0)}. \quad (31)$$

$P(y^s_{t\text{-}(n+1):t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1})$ in (31) is calculated using (22). The probability $P(r^s_{t\text{-}n\text{-}1}{=}s_1)$ is calculated using the following recursive relationship.

$$P(r_t^s){=}\Sigma_{r^s_{t\text{-}1}} P(r_t^s, r^s_{t\text{-}1}){=}\Sigma_{r^s_{t\text{-}1}} P(r_t^s|r^s_{t\text{-}1})P(r^s_{t\text{-}1}), \quad (32)$$

where $P(r_t^s|r^s_{t\text{-}1})$ is given by (3) and (4). $P(r_0^s{=}s_1){=}1$ and $P(r_0^s{=}s_0){=}0$ as we assume $r_0^s{=}s_1$. The probability $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r_t^s{=}s_0)$ is calculated as follows:

$$P(r^s_{t\text{-}(n+1):t\text{-}1}|r_t^s){=}P(r_t^s|r^s_{t\text{-}(n+1):t\text{-}1})\frac{P(r^s_{t\text{-}(n+1):t\text{-}1})}{P(r_t^s)} {=}P(r_t^s|r^s_{t\text{-}1})\frac{P(r^s_{t\text{-}(n+1)})}{P(r_t^s)}\left(\prod_{l=1}^{n} P(r^s_{t\text{-}l}|r^s_{t\text{-}l\text{-}1})\right). \quad (33)$$

This completes the calculations for Case 2.

If $t{=}0$, we have $r_0^s{=}s_1$ and

$$a_0^s{=}P(y_0^s, r_0^s{=}s_1){=}P(y_0^s|r_0^s{=}s_1), \quad (34)$$

which is calculated by (22). This completes the calculation of a_t^s.

Calculation of b_t^s: We assign 1 to b^s_{Ls}.

When $t \le Ls$-1, b_t^s in (19) is calculated as follows:

$$b_t^s{=}\Sigma_{n=0}^{t\text{-}1} P(y^s_{t+1:L_s}, r_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r_t^s) {=}\Sigma_{n=0}^{t\text{-}1} P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r_t^s) \times P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r_t^s) {=}\Sigma_{n=0}^{t\text{-}1} b^s_{t,n} P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r_t^s), \quad (35)$$

where $b^s_{t,n}(r_t^s){=}P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r_t^s)$. In the following, we describe the calculation of $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r_t^s)$ and $b^s_{t,n}(r_t^s)$.

The first probability $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r^s_t)$ is calculated as follows,

$$P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r^s_t) = \frac{P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r^s_t)}{P(y^s_{0:t}|r^s_t)} P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t). \quad (36)$$

Similar to the calculation of (27), the probability $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|y^s_{0:t}, r^s_t)$ is calculated based on the two probabilities $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r^s_t)$ and $P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t)$. $P(r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}|r^s_t)$ is calculated using (33). Now we calculate $P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t)$. We consider two cases. If $r^s_t{=}s_0$, the probability $P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t)$ is calculated using (22) by setting $n{=}t\text{-}1$. If $r^s_t{=}s_1$, we have

$$\begin{aligned} &P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t) \\ &{=}P(y^s_{0:t\text{-}1}|y^s_t, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t{=}s_1) \\ &\times P(y^s_t|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t{=}s_1) \\ &{=}P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1})P(y^s_t|r^s_t{=}s_1) \end{aligned} \quad (37)$$

$P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1})$ is calculated by (29) and $P(y^s_t|r^s_t{=}s_1)$ is calculated using (22).

The second probability $b^s_{t,n}(r^s_t)$ can be expressed as:

$$b^s_{t,n}(r^s_t){=}\Sigma_{r^s_{t+1}} P(y^s_{t+1:L_s}, r^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t). \quad (38)$$

If $t \leq L_s$-2, (38) is calculated using the following recursion:

$$\begin{aligned} b^s_{t,n}(r^s_t) &{=}\Sigma_{r^s_{t+1}} P(y^s_{t+2:L_s}|y^s_{0:t+1}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_{t:t+1}) \\ &\times P(r^s_{t+1}|r^s_t)P(y^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_{t:t+1}) \\ &{=}\begin{cases} \Sigma_{r^s_{t+1}} P(y^s_{t+1}|y^s_t, r^s_t{=}s_1, r^s_{t+1}) \\ \times b^s_{t+1,0} P(r^s_{t+1}|r^s_t), \text{if } r^s_t{=}s_1. \\ \Sigma_{r^s_{t+1}} P(y^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t}{=}S_{n+2}, r^s_{t+1}) \\ \times b^s_{t+1,n+1} P(r^s_{t+1}|r^s_t), \text{if } r^s_t{=}s_0. \end{cases} \end{aligned} \quad (39)$$

If $r^s_{t+1}{=}s_1$, $P(y^s_{t+1}|y^s_t, r^s_t{=}s_1, r^s_{t+1})$ and $P(y^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t}{=}S_{n+2}, r^s_{t+1})$ is calculated using (22). If $r^s_{t+1}{=}s_0$, both probabilities are calculated using (26). If $t{=}L_s$-1, (38) is expressed as:

$$\begin{aligned} b^s_{t,n}(r^s_t) &{=}\Sigma_{r^s_{t+1}} P(y^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_{t:t+1}) \\ &\times P(r^s_{t+1}|r^s_t) \\ &{=}\begin{cases} \Sigma_{r^s_{t+1}} P(y^s_{t+1}|y^s_t, r^s_t{=}s_1, r^s_{t+1}) \\ \times P(r^s_{t+1}|r^s_t), \text{if } r^s_t{=}s_1. \\ \Sigma_{r^s_{t+1}} P(y^s_{t+1}|y^s_{0:t}, r^s_{t\text{-}(n+1):t}{=}S_{n+2}, r^s_{t+1}) \\ \times P(r^s_{t+1}|r^s_t), \text{if } r^s_t{=}s_0. \end{cases} \end{aligned} \quad (40)$$

The probabilities in (40) are calculated in the same way as those in (39). This completes the calculation of b^s_t.

Calculation of $P(r^s_t|Y^s)$ in (14):

$$P(r^s_t|Y^s){=}\frac{P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_t)P(y^s_{0:t}, r^s_t)}{P(Y^s)} {=}\frac{a^s_t(r^s_t)b^s_t(r^s_t)}{P(Y^s)}, \quad (41)$$

where

$$P(Y^s){=}\Sigma_{r^s_t} a^s_t(r^s_t)b^s_t(r^s_t).$$

Similar to the calculation of (27), the probability $P(r^s_t|Y^s)$ is calculated using a^s_t, b^s_t.

Calculation of $P(r^s_{t\text{-}1}, r^s_t|Y^s)$ in (16):

$$\begin{aligned} &P(r^s_{t\text{-}1}, r^s_t|Y^s) \\ &{=}\frac{P(y^s_t, r^s_t, y^s_{t+1:L_s}|y^s_{0:t\text{-}1}, r^s_{t\text{-}1})P(y^s_{0:t\text{-}1}, r^s_{t\text{-}1})}{P(Y^s)} \\ &{=}\frac{P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}, r^s_t)P(y^s_t|y^s_{0:t\text{-}1}, r^s_{t\text{-}1}, r^s_t)}{P(Y^s)} \\ &\times P(r^s_t|r^s_{t\text{-}1})a^s_{t\text{-}1}. \end{aligned} \quad (42)$$

Similar to the calculation of probability in (41), we only need to caculate the probabilities $P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}, r^s_t), P(y^s_t|y^s_{0:t\text{-}1}, r^s_{t\text{-}1}, r^s_t), P(r^s_t|r^s_{t\text{-}1})$ and $a^s_{t\text{-}1}$.

In the following, we calculate $P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}, r^s_t)$ and $P(y^s_t|y^s_{0:t\text{-}1}, r^s_{t\text{-}1}, r^s_t)$ in (42). To calculate $P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}, r^s_t)$, we consider the following two cases:

Case 1: $r^s_{t\text{-}1}{=}s_1$,

$$P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}{=}s_1, r^s_t){=}b^s_{t,0}. \quad (43)$$

Case 2: $r^s_{t\text{-}1}{=}s_0$,

$$\begin{aligned} &P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t) \\ &{=}\Sigma^{t\text{-}1}_{n=1} P(y^s_{t+1:L_s}, r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t) \\ &{=}\Sigma^{t\text{-}1}_{n=1} P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t) \\ &\times P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t) \\ &{=}\Sigma^{t\text{-}1}_{n=1} b^s_{t,n} P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t), \end{aligned} \quad (44)$$

where

$$\begin{aligned} &P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t) \\ &{=}\frac{P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t)}{P(y^s_{0:t}|r^s_{t\text{-}1}{=}s_0, r^s_t)} \\ &\times P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|r^s_{t\text{-}1}{=}s_0). \end{aligned} \quad (45)$$

Similar to the calculation of (27), the probability $P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|y^s_{0:t}, r^s_{t\text{-}1}{=}s_0, r^s_t)$ is calculated by calculating the probabilities $P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|r^s_{t\text{-}1}{=}s_0)$ and $P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t)$. $P(r^s_{t\text{-}(n+1):t\text{-}2}{=}S_n|r^s_{t\text{-}1}{=}s_0)$ is calculated using (33). Based on the value of r^s_t, we have

$$\begin{aligned} &P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t) \\ &{=}\begin{cases} P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}) \\ \times P(y^s_t|r^s_t{=}s_1), \text{if } r^s_t{=}s_1. \\ P(y^s_{0:t}|r^s_{t\text{-}(n+1):t}{=}S_{n+2}), \text{if } r^s_t{=}s_0. \end{cases} \end{aligned} \quad (46)$$

If $r^s_t{=}s_1$,

$$\begin{aligned} &P(y^s_{0:t}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}, r^s_t) \\ &{=}P(y^s_t|r^s_t{=}s_1)P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1}), \end{aligned} \quad (47)$$

where $P(y^s_t|r^s_t{=}s_1)$ is calculated using (21), and $P(y^s_{0:t\text{-}1}|r^s_{t\text{-}(n+1):t\text{-}1}{=}S_{n+1})$ is calculated using (29). If $r^s_t{=}s_0$, the probability is calculated using (29). This completes the calculation of $P(y^s_{t+1:L_s}|y^s_{0:t}, r^s_{t\text{-}1}, r^s_t)$.

The probability $P(y_t^s|y_{0:t\text{-}1}^s, r_{t\text{-}1}^s, r_t^s)$ in (42) is calculated using (21) if $r_t^s{=}s_1$. The probability is calculated using (26) if $r_{t\text{-}1}^s{=}s_1$ and $r_t^s{=}s_0$. When $r_{t\text{-}1}^s{=}s_0$ and $r_t^s{=}s_0$, we have

$$\begin{aligned}&P(y_t^s|y_{0:t\text{-}1}^s, r_{t\text{-}1}^s{=}s_0, r_t^s{=}s_0)\\&{=}\Sigma_{n=1}^{t\text{-}1}P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|y_{0:t\text{-}1}^s, r_{t\text{-}1}^s{=}s_0, r_t^s{=}s_0)\\&\times P(y_t^s|y_{0:t\text{-}1}^s, r_{t\text{-}(n+1):t}^s{=}S_{n+2}),\end{aligned} \tag{48}$$

where

$$\begin{aligned}&P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|y_{0:t\text{-}1}^s, r_{t\text{-}1}^s{=}s_0, r_t^s{=}s_0)\\&{=}\frac{P(y_{0:t\text{-}1}^s|r_{t\text{-}(n+1):t}^s{=}S_{n+2})}{P(y_{0:t\text{-}1}^s|r_{t\text{-}1}^s{=}s_0, r_t^s{=}s_0)}\\&\times P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|r_{t\text{-}1}^s{=}s_0).\end{aligned} \tag{49}$$

Similar to the calculation of (27), the probability $P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|y_{0:t\text{-}1}^s, r_{t\text{-}1}^s{=}s_0, r_t^s{=}s_0)$ is calculated by calculating the probabilities $P(y_{0:t\text{-}1}^s|r_{t\text{-}(n+1):t}^s{=}S_{n+2})$ and $P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|r_{t\text{-}1}^s{=}s_0)$. The probability $P(y_t^s|y_{0:t\text{-}1}^s, r_{t\text{-}(n+1):t}^s{=}S_{n+2})$ in (48) and (49) is calculated using (26). $P(r_{t\text{-}(n+1):t\text{-}2}^s{=}S_n|r_{t\text{-}1}^s{=}s_0)$ is calculated with (33).

Calculation of $P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}|Y^s)$ in (14):

$$\begin{aligned}P(r_t^s{=}s_0|Y^s)&{=}\Sigma_{n=0}^{t\text{-}1}P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}|Y^s)\\&{=}\Sigma_{n=0}^{t\text{-}1}\frac{P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}, Y^s)}{P(Y^s)}.\end{aligned} \tag{50}$$

That is,

$$P(Y^s){=}\Sigma_{n=0}^{t\text{-}1}\frac{P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}, Y^s)}{P(r_t^s{=}s_0|Y^s)}. \tag{51}$$

Similar to the calculation of the probability in (27), the probability $P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}|Y^s)$ is calculated as

$$\frac{P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}, Y^s)}{P(r_t^s{=}s_0|Y^s)}, \tag{52}$$

where $P(r_t^s{=}s_0|Y^s)$ is given by (41), and

$$\begin{aligned}&P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}, Y^s)\\&{=}P(r_{t\text{-}(n+1):t}^s{=}S_{n+2}, y_{0:t\text{-}(n+2)}^s, y_{t\text{-}(n+1):t}^s, y_{t+1:L_s}^s)\\&{=}P(y_{t+1:L_s}^s|r_{t\text{-}(n+1):t}^s{=}S_{n+2}, y_{0:t}^s)\\&\times P(r_{t\text{-}(n+1):t}^s{=}S_{t\text{-}(n+1):t}, y_{0:t\text{-}(n+2)}^s, y_{t\text{-}(n+1):t}^s)\\&{=}b_{t,n}^s(s_0)P(y_{0:t\text{-}(n+2)}^s|r_{t\text{-}(n+1)}^s{=}s_1)P(r_{t\text{-}(n+1):t}^s{=}S_{n+2})\\&\times P(y_{t\text{-}(n+1):t}^s|r_{t\text{-}(n+1):t}^s{=}S_{n+2}).\end{aligned} \tag{53}$$

$P(y_{0:t\text{-}n\text{-}2}^s|r_{t\text{-}n\text{-}1}^s{=}s_1)$ is calculated using (30). $P(y_{t\text{-}n\text{-}1:t}^s|r_{t\text{-}(n+1):t}^s{=}S_{n+2})$ is calculated using (22). $P(r_{t\text{-}(n+1):t}^s{=}S_{n+2})$ is calculated with (32).

The probability $P(r_t^s|r_{t\text{-}1}^s, \theta)$ in (16) can be expressed as a function of θ using (3) and (4). The probabilities $P(y_t^s|r_t^s{=}s_1, \theta)$ and $P(y_t^s|y_{0:t\text{-}1}^s, r_{t\text{-}(n+1):t}^s{=}S_{n+2}, \theta)$ in (14) can be expressed as a function of θ using (22). The closed-form functions for the two probabilities based on the integration of $f(z;\theta_1)$ are needed for the M-step.

B. M-step

In this step, the estimate of model coefficients θ is updated by solving the optimization problem in (11). If a solution can be expressed in a closed-form, the update of θ^l can be easily obtained. If a closed-form solution cannot be obtained, numerical methods for nonlinear optimization problems (e.g., generalized reduced gradient method, sequential quadratic method, and interior point method) can be used to find an optimal solution. For example, a natural gradient iterative method is utilized in [13], [14] to solve the optimization problem in M-step with linear constraints for the estimates of transition matrix. In addition, since the objective function of the optimization model is not necessarily convex, local optimal solutions may exist. Multiple initial solutions and different algorithms can be applied to avoid local optimal solutions [9].

REFERENCES

[1] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, 2002.

[2] Y. Wan, K. Namuduri, Y. Zhou, and S. Fu, "A smooth-turn mobility model for airborne networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3359–3370, Sept 2013.

[3] J. G. Markoulidakis, G. L. Lyberopoulos, D. F. Tsirkas, and E. D. Sykas, "Mobility modeling in third-generation mobile telecommunications systems," *IEEE Personal Communications*, vol. 4, no. 4, pp. 41–56, Aug 1997.

[4] J. Xie, Y. Wan, J. H. Kim, S. Fu, and K. Namuduri, "A survey and analysis of mobility models for airborne networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1221–1238, Third 2014.

[5] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," in *Proceedings of IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–13.

[6] Y. Li, D. Jin, Z. Wang, P. Hui, L. Zeng, and S. Chen, "A markov jump process model for urban vehicular mobility: Modeling and applications," *IEEE Transactions on Mobile Computing*, vol. 13, no. 9, pp. 1911–1926, Sept 2014.

[7] J. Xie, Y. Wan, B. Wang, S. Fu, K. Lu, and J. H. Kim, "A comprehensive 3-dimensional random mobility modeling framework for airborne networks," *IEEE Access*, pp. 1 – 15, 2018.

[8] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," in *Proceedings of IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications*, April 2006, pp. 1–13.

[9] Andrew Logothetis and V. Krishnamurthy, "Expectation maximization algorithms for map estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, vol. 47, no. 8, pp. 2139–2156, Aug 1999.

[10] F. Zheng, S. Derrode, and W. Pieczynski, "Parameter estimation in conditionally gaussian pairwise markov switching models and unsupervised smoothing," in *Proceeding of 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, Sept 2016, pp. 1–6.

[11] A. Doucet, A. Logothetis, and V. Krishnamurthy, "Stochastic sampling algorithms for state estimation of jump markov linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 2, pp. 188–202, Feb 2000.

[12] E. zkan, F. Lindsten, C. Fritsche, and F. Gustafsson, "Recursive maximum likelihood identification of jump markov nonlinear systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 754–765, Feb 2015.

[13] M. Baro, J. S. Marques, and J. M. Lemos, "An improved em-method for the estimation of transition probabilities in multiple model switching systems," in *Proceedings of 8th IFAC Symposium on Nonlinear Control Systems*, vol. 43, no. 14, July 2010, pp. 374 – 378.

[14] J. S. M. M. Baro, "Offline bayesian identification of jump markov nonlinear systems," in *Proceedings of the 18th IFAC World Congress,Milan*, Aug 2011.