Appliance Load Disaggregation based on Bayesian Sequence Estimation using Importance Sampling

Venkata Pathuri-Bhuvana
Silicon Austria Labs & JKU LIT SAL eSPML Lab, Linz, Austria.
Email: venkata.pathuri-bhuvana@silicon-austria.com

***Abstract*—Appliance load disaggregation (ALD): Disaggregating the total power demand of an household into power demands of individual appliances can provide valuable information to consumers. In this paper, a non-intrusive ALD method based on Bayesian sequence estimation (BSE) is presented. Given a sequence of measured aggregated power demands of the household, the proposed method estimates the maximum probable state sequence of each appliance in the household by utilizing the Viterbi framework. In addition, this paper proposes three importance sampling (IS) methods namely uniform, transition and weighted transition methods to reduce the complexity of the proposed BSE based ALD by state reduction. Furthermore, the proposed method has a backtracing mechanism to smooth the estimated state sequence based on the future measurements. The proposed method is compared with the particle filter (PF) based ALD using two well known real-world data sets. The simulation results show that the proposed method achieves higher estimation accuracy than the PF based ALD method depending on the chosen IS method.**

I. Introduction

Modern smart meters enable the consumers to monitor fine-grained power demand data of their households in real time [1]. Disaggregating the overall power demand of an household into the power demands of individual appliances, often referred as appliance load disaggregation (ALD) can enhance the energy aware behavior of the consumer. Several authors have proposed both intrusive and non-intrusive ALD methods [2] based on machine learning approaches [3]–[5], matrix factorization [6], integer programming [7], [8] and hidden markov model (HMM) and factorized hidden markov model (FHMM) [9], [10], and so on. In [8] and [9], it is shown that the HMM-FHMM based ALD methods have been preferred over other approaches. The HMM-FHMM based ALD methods utilize the probabilistic appliance models to model a household as a finite state machine. They can further use highly scalable particle filters (PF) to estimate the appliance states based on maximum a-posterior (MAP) criteria [9].

The MAP based ALD method proposed in [9] uses a relatively efficient reduced state Viterbi algorithm. However, this method estimates the state of each appliance at a given time based on the aggregated power demand of the household at that time. Thus, it does not exploit the information carried by the future aggregated power demand measurements of the household. Further, it uses random sampling to achieve state reduction without considering the appliance behavior or state transition probabilities.

In this paper, an ALD method based on Bayesian sequence estimation (BSE) with important sampling (IS) and backtracing is proposed. Contrary to [9], the proposed method estimates the sequence of appliance states in a given observation window based on all the aggregated power demand measurements in that observation window. The proposed method uses the Viterbi algorithm and IS to calculate the maximum probable state sequence of each appliance by maximizing the joint probability density function (PDF) of all the appliance states and aggregated power demand measurements in a given observation window. The role of IS is to reduce the computational complexity of the BSE by state reduction thus, resulting in a reduced state BSE (RBSE) based ALD. This paper proposes three different IS methods, namely uniform (UIS), transition (TIS) and weighted transition (WIS) to allow the proposed RBSE based ALD method to perform state reduction. The proposed method further has a backtracing mechanism that requires the estimation process to come backwards and smooth the estimated state sequence based on the future measurements. Furthermore, the proposed method is evaluated using two real world data sets, namely a public data set for energy disaggregation research (REDD) [11] and an energy consumption data set of households in Italy and Austria (GREEND) [12]. For the considered data sets, the proposed RBSE based ALD method with IS and backtracing can achieve higher accuracy compared to the PF based ALD method.

The paper is organized as follows, Sec. II describes the HMM and FHMM based appliance and household models, respectively. Sec. III describes the BSE based ALD. Sec. IV describes the proposed RBSE based ALD algorithm along with the proposed IS methods and backtracing process. Sec. V evaluates the proposed method based on REDD and GREEND data sets. Sec. VI presents the conclusions.

II. System Model

Let us consider that the household consists of N appliances and that an appliance n, where $n = 1, 2, \cdots, N$ is operating in a state $x_k^n \in \{s_1^n, s_2^n, \cdots, s_{S^n}^n\}$ at time k. The variable S^n represents the number of possible states in which the appliance n can be operated. The behavior of the appliance over time can be modeled using the morkov model (MM) as shown in Fig. 1. The MM is characterized by the state transition matrix $\mathbf{A}_k^n$ of the appliance n at time k. The element $a_{i,j,k}^n$, where $1 \leq i, j \leq S^n$ in the state transition matrix $\mathbf{A}_k^n$ represents the transition probability of the appliance n from state s_i^n to state s_j^n at time k as

$$a_{i,j,k}^n = \Pr\left(x_k^n = s_j^n | x_{k-1}^n = s_i^n\right) \geq 0, \tag{1}$$

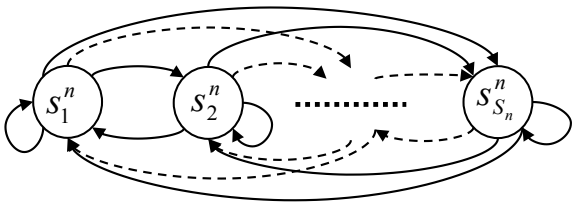

Fig. 1. The probabilistic appliance model based on the markov model.

where $\sum_{j=1}^{S_n} a_{i,j,k}^n = 1 \; \forall \; i,k$. The elements $a_{i,j,1}^n$ in the initial state transition matrix $\mathbf{A}_1^n$ are defined as

$$a_{i,j,1}^n = \Pr\left(x_1^n = s_j^n | x_0^n\right) = \pi_j^n \geq 0, \tag{2}$$

where x_0^n is the initial state of the appliance n.

The power demand $P_k^n \in \left\{P(s_1^n), P(s_2^n), \cdots, P(s_{S_n}^n)\right\}$ of the appliance n at time k corresponds to the state x_k^n in which the appliance n is operated at the time k. The overall household power demand P_k at time k is the aggregate of the power demands of N household appliances at time k and it can be modeled by FHMM. The FHMM models the overall household power demand as a superposed power demand of N multiple Markov chains, where each Markov chain represents an individual appliance. Thus, the measured aggregated power P_k is given by

$$P_k = \sum_{n=1}^{N} P_k^n + w_k, \tag{3}$$

where the measurement error w_k is modeled with an independent and identically distributed (IID) Gaussian noise.

III. Appliance Load Disaggregation using Bayesian Sequence Estimation

Let us consider that $\mathbf{x}_k = \left[x_k^1, x_k^2, \cdots, x_k^N\right]^T$ represents the operating states of all N appliances (super state of the household) at time k. Given the aggregated power demands of an household $P_{1:K} = (P_1, P_2, \cdots, P_K)$ in an observation window, the goal is to estimate the sequence of maximum probable super states $\mathbf{x}_{1:K} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_K)$. The estimated state sequence $\hat{\mathbf{x}}_{1:K}$ can be inferred by maximizing the joint PDF $p(\mathbf{x}_{1:K}, P_{1:K})$ [13] as

$$\hat{\mathbf{x}}_{1:K} = \arg\max_{\mathbf{x}_{1:K}} p(\mathbf{x}_{1:K}, P_{1:K}). \tag{4}$$

According to the assumed appliance model in (1) and household model in (3), the super state $\mathbf{x}_K$ and the power demand P_K of the household only depend on the super state $\mathbf{x}_{K-1}$ and $\mathbf{x}_K$, respectively. Thus, the joint PDF in (4) can be written as

$$p(\mathbf{x}_{1:K}, P_{1:K}) = p(P_K|\mathbf{x}_K)p(\mathbf{x}_K|\mathbf{x}_{K-1})p(\mathbf{x}_{1:K-1}, P_{1:K-1}), \tag{5}$$

where $p(P_K|\mathbf{x}_K)$ and $p(\mathbf{x}_K|\mathbf{x}_{K-1})$ are the likelihood and state transition PDFs at time K, respectively. $p(\mathbf{x}_{1:K-1}, P_{1:K-1})$ is the joint PDF at time $K-1$.

The Viterbi based BSE (VBSE) can be used to maximize the joint PDF as shown in (4). According to the assumed HMM based appliance model in (1), the household can reach a certain super state $\mathbf{x}_k^i$ at each time k, where $i = 1, 2, \cdots, S$ and $k \in 1, 2, \cdots, K$. The variable $S = \prod_{n=1}^{N} S^n$ represents the total number of possible super states of the household, where S^n is the number of possible states of the appliance n. The VBSE computes maximum likelihood (ML) measure α_k^i for each super state $\mathbf{x}_k^i$ at each time k. The ML measure α_k^i gives the maximum probability for the household to reach the super state $\mathbf{x}_k^i$ at time k and it can be written as

$$\begin{aligned}
\alpha_k^i &= \max_{\mathbf{x}_{1:k-1}} \ln p(\mathbf{x}_k = \mathbf{x}_k^i, \mathbf{x}_{1:k-1}, P_{1:k}), \\
&= \ln p(P_k|\mathbf{x}_k = \mathbf{x}_k^i) + \max_j \big[\ln p(\mathbf{x}_k = \mathbf{x}_k^i | \\
&\qquad \mathbf{x}_{k-1} = \mathbf{x}_{k-1}^j) + \ln p(\mathbf{x}_{1:k-1}|P_{1:k-1})\big], \\
&= \ln p(P_k|\mathbf{x}_k = \mathbf{x}_k^i) + \max_j \big[\ln p(\mathbf{x}_k = \mathbf{x}_k^i | \\
&\qquad \mathbf{x}_{k-1} = \mathbf{x}_{k-1}^j) + \alpha_{k-1}^j\big],
\end{aligned} \tag{6}$$

where α_{k-1}^j is the ML measure to reach the state $\mathbf{x}_{k-1}^j$ at time $k-1$, where $j = 1, 2, \cdots, S$. The initial ML measure α_1^i at time $k = 1$ and for $i = 1, 2, \cdots, S$ is defined as

$$\alpha_1^i = \ln p(P_1|\mathbf{x}_1 = \mathbf{x}_1^i) + \ln p(\mathbf{x}_1 = \mathbf{x}_1^i|\mathbf{x}_0). \tag{7}$$

Further, the VBSE identifies the path ψ_k^i that maximizes the ML measure α_k^i and it is given as

$$\psi_k^i = \arg\max_j \left[\ln p(\mathbf{x}_k = \mathbf{x}_k^i|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^j) + \alpha_{k-1}^j\right]. \tag{8}$$

In order to optimally estimate the maximum probable state sequence $\hat{\mathbf{x}}_{1:K}$, the VBSE should calculate the ML measures $\alpha_k^i|_{i=1}^S|_{k=1}^K$ and the corresponding paths $\psi_k^i|_{i=1}^S|_{k=1}^K$ for all possible super states at each time k until K. This is refereed as forward recursion. At end of the forward recursion, the optimal sequence of super states can be estimated using a process known as backtracing. Firstly, the estimated super state $\hat{\mathbf{x}}_K$ of the household is computed as

$$\hat{\mathbf{x}}_K = \mathbf{x}_K^{i_K} \quad \text{where} \quad i_K = \arg\max_i \alpha_K^i, \tag{9}$$

Now, the backtracing process keeps the track of the paths ψ_k^i that maximizes ML measures α_k^i for each time $k \in K-1, K-2, \cdots, 1$. By applying backtracing, the super state $\mathbf{x}_k$ at time k, where $k = K-1, K-2, \cdots, 1$ can be estimated as

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^{i_k} \quad \text{where} \quad i_k = \psi_{k+1}(i_{k+1}). \tag{10}$$

The estimated state sequence $\hat{\mathbf{x}}_{1:K}$ is the optimal state sequence of the household from time 1 to K given the power demand measurements $P_{1:K}$. Assuming that the household has S operating super states, the computational complexity and the memory requirements of the optimal VBSE method are in order of $O(S^2K)$ and $O(2SK)$, respectively. Hence, in practice, the implementation of the optimal BSE to solve the ALD problem could already become computationally intensive for an household with as many as 10 ON-OFF appliances. In the next section, a reduced state BSE is proposed which adopts the Viterbi paradigm as discussed in this section and IS methods to estimate the appliance states.

IV. The Reduced State Bayesian Sequence Estimation for Appliance Load Disaggregation

In the proposed RBSE method, the number of super states at which the ML measures α_k^i and the corresponding maximizing paths ψ_k^i are calculated is restricted to a sub set of super states. In this paper, a monte-carlo process known as importance sampling (IS) [14] is used to deduce the subset of super states at which the ML measures α_k^i and the corresponding maximizing paths ψ_k^i are calculated.

Let us consider that at time k, a set of particles $\mathbf{x}_{1:k-1}^{(j)}$, where $j = 1, 2, \cdots, N_p$ and the corresponding ML measures $\alpha_{k-1}^{(j)}$ are known. Each of these particles represent a super state of the household at time $k-1$. Now, a new set of particles $\mathbf{x}_{1:k}^{(i)}$, where $i = 1, 2, \cdots, N_p$ are drawn such that the corresponding ML measures $\alpha_k^{(i)}$ at time k can be calculated as required by the VBSE. Each of the particles in the new set $\mathbf{x}_{1:k}^{(i)}$ represent a super state of the household at time k. Ideally, the particles $\mathbf{x}_{1:k}^{(i)}$ should be drawn from the joint PDF $p(\mathbf{x}_{1:k}, P_{1:k})$ itself or any other PDF that has the same global maximum and support space as $p(\mathbf{x}_{1:k}, P_{1:k})$. Direct sampling from $p(\mathbf{x}_{1:k}, P_{1:k})$ is usually not possible. Hence, an importance density $q(\mathbf{x}_{1:k})$ is used to draw the samples at time k. The accuracy of the estimation depends on the selection of the importance density and the number of particles N_p.

Three different IS methods that use three different importance densities are discussed later in the paper. At this point, let us consider that the new set of particles $\mathbf{x}_{1:k}^{(i)}$, where $i = 1, 2, \cdots, N_p$ are drawn from one of these importance densities $q(\mathbf{x}_{1:k})$. The ML measure $\alpha_k^{(i)}$ at time k gives the maximum probability for the household to reach the super state (particle) $\mathbf{x}_k^{(i)}$, where $i = 1, 2, \cdots, N_p$. From (6), it can be written as

$$\alpha_k^{(i)} = \ln p(P_k|\mathbf{x}_k = \mathbf{x}_k^{(i)}) + \max_j \big[\alpha_{k-1}^{(j)} + \ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)})\big]. \tag{11}$$

The path $\psi_k^{(i)}$ that maximizes ML measure $\alpha_k^{(i)}$ is given as

$$\psi_k^{(i)} = \arg\max_j \left[\ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)}) + \alpha_{k-1}^{(j)}\right], \tag{12}$$

where the likelihood probability $p(P_k|\mathbf{x}_k = \mathbf{x}_k^{(j)})$ can be obtained from (3). The transition probability $p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)})$ can be calculated using the chosen importance density.

A. Importance Sampling

1) Uniform Importance Sampling (UIS): A new set of samples $\mathbf{x}_{1:k}^{(i)}$ are drawn uniformly among all the possible super states $\mathbf{x}_k^j$, where $j = 1, 2, \cdots, S$. Hence, the log of the transition probability $\ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)})$ that is used to compute the ML variable $\alpha_k^{(i)}$ as shown in (11) becomes

$$\ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)}) = \ln 1/S. \tag{13}$$

The UIS does not take the state transition probability $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ into account.

2) Transition Importance Sampling (TIS): In this approach, the new set of particle $\mathbf{x}_{1:k}^{(i)}$ at time k are drawn from the state transition PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. Since the particles $\mathbf{x}_{1:k-1}^{(j)}$, where $j = 1, 2, \cdots, N_p$ at time $k-1$ are known, the new set of particles $\mathbf{x}_{1:k}^{(i)}$, where $i = 1, 2, \cdots, N_p$ are drawn as

$$\mathbf{x}_{1:k}^{(i)}\Big|_{i=1}^{N_p} \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(j)})\Big|_{j=1}^{N_p}. \tag{14}$$

The TIS can make use of the HMM based appliance modeling to sample the new set of particles at each time k. Hence, the log of the transition probability that is used to compute the ML measure $\alpha_k^{(i)}$ and the maximizing path $\psi_k^{(i)}$ as shown in (11) and (12), respectively, is calculated as

$$\ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)}) = \sum_{n=1}^{N} \ln a_{i,j,k}^n. \tag{15}$$

This approach suffers from the degeneracy in which each particle $\mathbf{x}_{0:k-1}^{(i)}$ at time $k-1$ generates only one particle $\mathbf{x}_{0:k}^{(i)}$ at time k irrespective of its ML measure $\alpha_{k-1}^{(i)}$. Due to the this, large computational effort is spent on the particles with less contribution to the estimation of the optimal state sequence.

3) Weighted Importance Sampling (WIS): This approach partially addresses the degeneracy problem by allowing a particle $\mathbf{x}_{1:k-1}^{(j)}$ at time $k-1$ to generate one or more samples at time k based on its ML measure $\alpha_{k-1}^{(j)}$. In the WIS approach, the ML measure $\alpha_{k-1}^{(j)}$ of a sample $\mathbf{x}_{1:k-1}^{(j)}$ is normalized as

$$\overline{\alpha}_{k-1}^{(j)} = \alpha_{k-1}^{(j)} \Big/ \sum_{j=1}^{N_p} \alpha_{k-1}^{(j)}, \tag{16}$$

where $\overline{\alpha}_{k-1}^{(j)}$ is the normalized ML measure of the particle $\mathbf{x}_{1:k-1}^{(j)}$. Then, each particle $\mathbf{x}_{1:k-1}^{(j)}$, where $j = 1, 2, \cdots, N_p$ generates a certain number of particles proportional to the value $\overline{\alpha}_{k-1}^{(j)}$ using the corresponding state transition probability $p(\mathbf{x}_k|\mathbf{x}_{k-1}^{(i)})$ in such a way that total number of new particles remain N_p. Finally, the log of the transition probability $\ln p(\mathbf{x}_k = \mathbf{x}_k^{(i)}|\mathbf{x}_{k-1} = \mathbf{x}_{k-1}^{(j)})$ that is used to compute the ML variable $\alpha_k^{(i)}$ can be calculated as shown in (15).

In the RBSE approach the ML measures $\alpha_k^{(i)}\big|_{i=1}^{N_p}\big|_{k=1}^{K}$ and the corresponding maximizing paths $\psi_k^{(i)}\big|_{i=1}^{N_p}\big|_{k=1}^{K}$ are calculated for all sampled particles until time K. Then backtracing is used to achieve the estimated state sequence $\hat{\mathbf{x}}_{1:K}$ as described in the previous section. In general, the ALD algorithms run for a very long time (in the order of days), thus the RBSE based ALD with backtracing can cause high latency in estimating the state sequence. To reduce the latency to a required level, a truncated backtracing is proposed.

B. Truncated Backtracing

In this process, a latency factor L is introduced and backtracing is applied after every L power demand measurements. Let us consider that the total set of power demand measurements $P_{1:K}$ is divided into sub-blocks $\{P_{1:L}, P_{L+1:2L}, \cdots, P_{(K/L-1)L+1:K}\}$ of equal length L.

TABLE I

THE ACC OF THE PROPOSED RBSE BASED ALD WITH AND WITHOUT BACKTRACING FOR THE HOUSE 0 OF THE GREEND AND HOUSE 0 OF THE REDD DATA SETS IN COMPARISON TO PF BASED ALD METHOD. THE LATENCY FACTOR L IS CONSIDERED TO BE 1200. THE N_p IS 300.

Data set	Appliance	no. of states	RBSE-ALD no backtracing			RBSE-ALD backtracing			PF-ALD
			UIS	TIS	WIS	UIS	TIS	WIS	
GREEND	Coffee machine	3	0.5735	0.6431	0.6576	0.6341	0.7932	0.8021	0.6391
	Wash machine	4	0.7245	0.8546	0.8622	0.7242	0.9235	0.9634	0.8401
	Fridge	4	0.7323	0.8438	0.8460	0.7420	0.8845	0.8712	0.8278
	Dishwasher	3	0.8134	0.8536	0.8744	0.8203	0.8904	0.9132	0.8131
	Television	2	0.8451	0.8640	0.8715	0.8112	0.8833	0.9589	0.8536
Household	-	288	0.7377	0.8117	0.8224	0.7463	0.8749	0.9017	0.7951
REDD	Oven	3	0.7532	0.9324	0.9342	0.7832	0.9532	0.9534	0.9321
	Refrigerator	3	0.6234	0.7825	0.8023	0.7134	0.8412	0.8631	0.7843
	Dishwasher	4	0.5934	0.7722	0.7823	0.6345	0.8334	0.8505	0.7324
	Kitchen Outlet	2	0.8234	0.9374	0.9452	0.8134	0.9453	0.9464	0.9342
	Microwave	3	0.7134	0.8134	0.8232	0.7435	0.8634	0.8918	0.7924
Household	-	216	0.7013	0.8475	0.8574	0.7376	0.8877	0.9010	0.8350

When k is equal to mL, where $m = 1, 2, \cdots K/L$, the state sequence $\hat{\mathbf{x}}_{(m-1)L+1:mL}$ is estimated. By applying truncated backtracing, firstly, the estimated super state $\hat{\mathbf{x}}_{mL}$, where $m = 1, 2, \cdots K/L$ of the household is computed as

$$\widehat{\mathbf{x}}_{mL} = \mathbf{x}_{mL}^{i_{mL}} \quad \text{where} \quad i_{mL} = \arg\max_{i} \left\{ \alpha_{mL}^{(i)} \right\}. \tag{17}$$

Now, the super states $\mathbf{x}_k$ at time k, where $k = mL-1, mL-2, \cdots, mL-L+1$ can be estimated as

$$\widehat{\mathbf{x}}_k = \mathbf{x}_k^{i_k} \quad \text{where} \quad i_k = \psi_{k+1}^{(i_{k+1})}. \tag{18}$$

The estimated super state $\hat{\mathbf{x}}_{mL}$ is then used as the prior state for the estimation of the state sequence $\mathbf{x}_{mL+1:(m+1)L}$. An higher estimation accuracy can be achieved by either increasing the latency factor L or the number of particles N_p which also increases the latency of the estimation and/or memory and computational requirements. For a chosen L and N_p, the memory requirements and computational complexity are in order of $O(2N_p \cdot L)$ and $O(N_p^2 \cdot L)$, respectively.

V. RESULTS

In this section, the evaluation of the proposed RBSE based ALD is presented in comparison to the PF based ALD [9]. Two different real world data sets REDD and GREEND are used to evaluate the performance of the proposed method. For our evaluations, we used the data from House 1 in REDD and House 0 in GREEND data sets. The metrics accuracy and root mean square error (RMSE) are used for the performance evaluation. The accuracy is calculated based on true detection (TD) (number of times the state of an appliance is correctly detected) and false detection (FD) (number of times the state of an appliance is wrongly detected) as

$$\mathrm{ACC} = \mathrm{TD} / \left(\mathrm{TD} + \mathrm{FD}\right) \quad \in [0, 1]. \tag{19}$$

The normalized RMSE is formulated as

$$\mathrm{RMSE} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(P_k - \hat{P}_k\right)^2} \Big/ \left(P_{\max} - P_{\min}\right), \tag{20}$$

where P_k and $\hat{P}_k$ are the true power demand and the estimated power demand of the household at time k, respectively. $P_{\max}$ and $P_{\min}$ are the maximum and the minimum aggregated power demand of the household in the selected window of time, respectively. K is the duration of the estimation process. In order to be statistically reliable, the accuracy ACC and RMSE are averaged over 100 simulation runs.

The considered data sets REDD and GREEND provide the sub-metered power demands of the individual appliances. The total power demand of the household is determined by aggregating the power demands of the chosen individual appliances. The proposed RBSE and PF based ALD methods require the state transition matrix $\mathbf{A}_k^n$ for each appliance n in order to estimate the state of the appliances. In this paper, the transition matrix $\mathbf{A}_k^n$ for each device n is determined by using 30% of the power demand measurements. The Baum-Welch algorithm [15] is used to construct the state transition matrices $\mathbf{A}_k^n$ from the training data. Moreover, the state transition matrix $\mathbf{A}_k^n$ for each device n is assumed to be constant for the entire duration of the estimation process. Based on the power demand information, each appliance is modeled as either ON/OFF appliance or multi-state appliances, where we give the algorithm the possibility to adjust its power demand for each state. Furthermore, the Baum-Welch algorithm also calculates an observation matrix $\mathbf{B}_k^n$ which can be used to calculate the estimated power $\hat{P}_k^n$ of each device n based on the estimated state $\hat{x}_k^n$. Thus, the estimated power demand $\hat{P}_k$ of the household can be calculated as an aggregate of the estimated power $\hat{P}_k^n$ of individual appliances.

Table I gives the accuracy ACC of the proposed RBSE based ALD method with and without backtracing in comparison with the PF based ALD. For this evaluation, the number of particles N_p and latency factor L are considered to be 300 and 1200, respectively. When backtracing is not applied, the proposed method estimates the appliance state at a given time based on the power demand measurement at that time similar to the PF-MAP based ALD proposed in [15]. From

TABLE II
THE RMSE OF THE PROPOSED RBSE BASED ALD WITH BACKTRACING FOR THE HOUSE 0 OF THE GREEND DATA SET.

No. of particles N_p	RBSE with backtracing		
	UIS	TIS	WIS
$N_p = 100$	0.1732	0.1256	0.1221
$N_p = 300$	0.1627	0.1145	0.1086

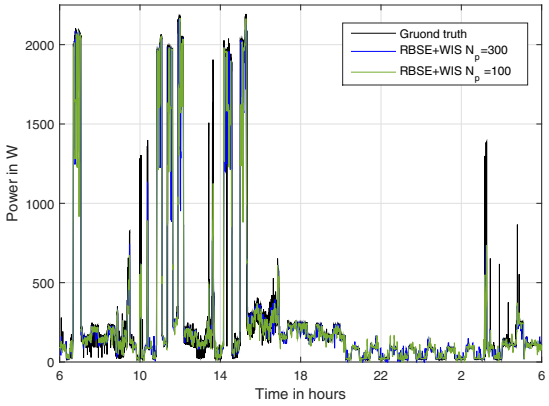

Fig. 2. A snap shot of the ground truth and estimated total power demands of the Household 0 of the GREEND data set. The estimation method is RBSE with backtracing and WIS. The number of particles N_p are considered to be 100 and 300. The latency factor L is 1200.

this table, it is illustrated that the accuracy of the proposed RBSE based ALD method increases when the backtracing is applied irrespective of the IS method. In general, it can also understood that the IS methods TIS and WIS are more accurate than the UIS method due to their ability use the state transition probability to sample the new particles instead of random sampling. The proposed RBSE based ALD method with UIS is also less accurate than the PF based ALD due to the same reason. On the other hand, the proposed RBSE base ALD method with TIS and WIS are more accurate than the PF based ALD method.

Table II gives the RMSE of the proposed RBSE based ALD with backtracing for the House 0 of the GREEND data set for all three IS methods. For this evaluation, two different number of particles N_p 100 and 300 are considered. The latency factor L is considered to be 1200. The RMSE mentioned in the table II corresponds to the entire household instead of individual appliances. From table II, it can be understood that the proposed RBSE method with WIS has better RMSE compared to that of the UIS and TIS methods. Further, Fig. 2 illustrates a snap shot of the ground truth and the estimated total power demands of the Household 0 of the GREEND data set for a chosen 24 hour period and in a simulation run using RBSE based ALD with backtracing and WIS. This figure shows that the RMSE of the proposed method decreases as the number of particles N_p increases. Moreover, the proposed methods has better estimation accuracy when the power demand has less fluctuations.

VI. CONCLUSIONS

In this paper, we have proposed a reduced state Bayesian sequence estimation (RBSE) approach to perform appliance load disaggregation (ALD). The proposed method uses the principles of the importance sampling (IS) and the Viterbi algorithm to estimate the maximum probable state sequence of the appliances given the corresponding sequence of aggregated power demand measurements of the household. Three different IS methods: uniform, transition and weighted transition have been proposed. Unlike the particle filter (PF) based ALD, the proposed RBSE method performs backtracing to come backwards and smooth the estimated state sequence based on the future measurements. To further reduce the memory requirements and latency of the disaggregation process, a truncated backtracing has been proposed. The proposed method has been evaluated using REDD and GREEND in comparison with the PF based ALD. For the considered data sets, the proposed RBSE based ALD method can achieve higher accuracy compared to the PF based ALD method.

ACKNOWLEDGMENTS

This work has also been supported by the "University SAL Labs" initiative of Silicon Austria Labs (SAL) and its Austrian partner universities for applied fundamental research for electronic based systems.

REFERENCES

[1] L. Peretto, "The Role of Measurements in the Smart Grid Era," in *IEEE Instrumentation & Measurements Magazine*, vol. 13, no. 3, pp. 22-25, 2010.
[2] Z. Ahmed et al., "Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey"', in *Sensors*, vol. 12, no. 12, pp. 16838-16866, 2012.
[3] M. A. Mengistu et al., "A Cloud-Based On-Line Disaggregation Algorithm for Home Appliance Loads", in *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3430-3439, 2019.
[4] M. Xia et al., "Non-Intrusive Load Disaggregation based on Deep Dilated Residual Network", in *Electric Power Systems Research*, vol. 170, pp. 277-285, 2019.
[5] C. Nalmpantis et al., "Machine Learning Approaches for Non-Intrusive Load Monitoring: from Qualitative to Quantitative Comparison", in *Artificial Intelligence Review*, vol. 52, pp. 217–243, 2019.
[6] A. Miyasawa et al., "Energy Disaggregation based on Smart Metering Data via Semi-Binary Non-negative Matrix Factorization", in *Energy and Buildings*, vol. 183, pp. 547-558, 2019.
[7] M. Z. A. Bhotto et al., "Load Disaggregation Based on Aided Linear Integer Programming", in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, no. 7, pp. 792-796, 2017.
[8] W. Kong et al., "Improving Nonintrusive Load Monitoring Efficiency via a Hybrid Programing Method", in *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2148-2157, 2016.
[9] D. Egarter et al., "PALDi: Online Load Disaggregation via Particle Filtering," in *IEEE Transactions on Instrumentation & Measurement* , vol. 64, no. 2, pp. 467-477, 2015.
[10] V. P. Bhuvana et al., "Appliance Load Disaggregation based on Maximum A-Posterior Particle Filter (MAP-PF)," In *proceedings of IEEE Conference on Smart Grid Communications*, pp. 165-170, 2017.
[11] J. Z. Kolter et al., "REDD: A Public Data Set for Energy Disaggregation Research," in *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, 2011.
[12] A. Monacchi et al., "GREEND: An Energy Consumption Dataset of Households in Italy and Austria," In *arXiv:1405.3100*, 2014.
[13] S. M. Kay, "Fundamentals of Statistical Signal Processing:Estimation Theory," *Prentice Hall*, New Jersey, USA, vol. 1, 1993.
[14] M. S. Arulampalam et al., "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," in *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174-188, 2002.
[15] R. Durbin et al., "Biological Sequence Analysis," in *Cambridge University Press*, Cambridge, UK, 1998.