A Maximum-Entropy Based Fast Estimation of Power Quality for Smart Microgrid

Sardar Ali, Kui Wu, and Dimitri Marinakis
Department of Computer Science
University of Victoria, Victoria, BC, Canada

Abstract—**Power quality plays a critical role in the reliability of power networks. The monitoring of power quality, however, is a non-trivial task due to the high cost of measurement devices and the requirement on real-time responses. As a result, we need fast algorithms to estimate the power quality in various segments of a power network, using only a small number of measurement points. In this paper, we propose a maximum-entropy (MaxEnt) based approach to estimating power quality in smart microgrid. Compared to other existing methods such as Monte Carlo Expectation Maximization (MCEM), the MaxEnt based approach is much faster.**

I. INTRODUCTION

Power quality largely impacts the reliability and energy saving of power networks. Poor power quality such as voltage sags may lead to power outage and service interruptions. Service unavailability caused by power losses is a serious problem for many companies and organizations, e.g., it may result in a significant revenue loss for Internet service providers or even loss of lives in hospitals. To improve the reliability of power networks, organizations and large companies (e.g., Google data centers) adopt smart microgrid, and closely monitor the power quality in different segments of the microgrid.

Monitoring power quality, however, is not an easy task. Since the power measurement devices [1][2] (termed as smart meters in this paper) are expensive, it is financially impractical to monitor every segment of a power network. The overhead of interconnecting these power meters and developing the power management system further increases the cost. In addition, in many cases direct monitoring of power quality is difficult, e.g., it is hard to install smart meters after power lines were sealed in hard-to-reach areas in a building. In general, we need to tackle the following challenge: based on a limited small number of monitored points in a power network, how can we effectively estimate the power quality of other unmonitored segments of the network?

In this paper, first we model the smart microgid as a data-driven network where we represent the electrical components as network nodes, power links as data links, flow of power as data flow on the links, and the power flowing through links as numeric data. We then model the power quality estimation problem as an optimization problem of missing data estimation in a data network.

We propose to use a Maximum-Entropy (MaxEnt) [3] approach to power quality estimation. The basic idea of MaxEnt is that out of all probability distributions consistent with a given set of constraints (i.e., the known measure values in our case), we should choose the one that has the maximum uncertainty to be the estimated power quality values. Intuitively, the principle of MaxEnt implies that we should make use of all the information that is given and avoid making (biased) assumptions about information that is not available. We model the problem of estimating power quality in such a way where we can effectively get the benefit of MaxEnt approach to correctly estimate the power quality values at links where we do not have any measuring device installed. We solve the formulated MaxEnt problem and validate its effectiveness and efficiency with a simulated microgrid system.

The rest of the paper is organized as follows. The related work is summarized in Section II. To facilitate understanding and problem illustration, we (analogically) model a microgrid as a data-driven network in Section III. In Section IV, we formulate the power quality estimation problem as a MaxEnt problem. In Section V, we test the performance of MaxEnt solution with respect to its running speed and estimation errors, using a simulated microgrid. We also compare its performance to that of Expectation Maximization (EM) [4] based algorithm. We conclude the paper in Section VI.

II. RELATED WORK

According to the IEEE Gold Book [5], the industry practices for electric power quality in networks focus on measures such as reliability, availability, and Mean Time Between Failures (MTBF). While it is known [6] that lifetime and performance of electrical components are dependent on power quality, none of the measures defined in the IEEE Gold Book accounts for power quality. Moreover, these measures serve as theoretical values for comparisons and are not intended to work in real environment (with varying temperature, humidity, load, and power quality) as predictive tools for networks. Further, the ITI curve report illustrates the relationship between power quality and the likelihood of damage to electric components. While considering the importance of both the magnitude and duration of power quality events in isolation, it does not consider its cumulative effects over time.

The EM algorithm is one of the most widely-used algorithms for estimation and has been applied in a variety of research areas. In [7], the authors investigate the power of EM algorithm in the estimation of microgrid reliability. Although effective, the EM algorithm needs a long time to converge. We are thus motivated to find faster algorithms in power quality estimation for smart microgrid.
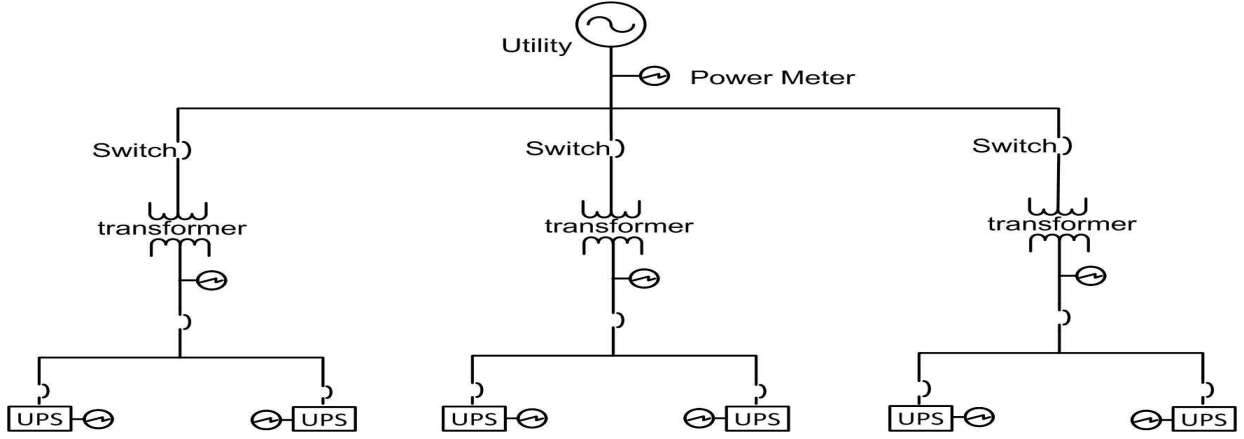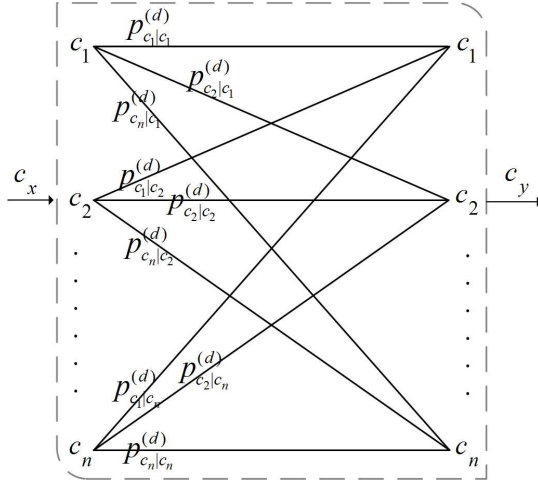
Fig. 1. A simple view of power microgrid

Fig. 2. Power quality transition at each device d as a channel.

Fig. 3. View of the power grid network under consideration. Subnets (within dotted lines) are formulated based on the positions of the power meters.

The measurement and monitoring of power quality are a vital part of today's smart grid. To classify power quality, a label is assigned to a power quality event, based on the feature of the event, e.g., the magnitude and duration of a voltage sage or swell. Typically, the power quality index, also known as System Average RMS Variation Frequency Index (SARFI), is reported as a count of the number of times the magnitude and duration fall below a threshold, standardised by the IEEE standard 1159-2009 [8].

III. MICROGRID AS A DATA DRIVEN NETWORK

We now model a smart microgrid as a data driven network, *analogically*, where electrical components are represented as nodes, power links as data links, and power flow as data flow. Hence the power grid can be viewed as a data network where we have nodes connected via data links and the data links transmit numeric data (i.e., power quality values/events in our case).

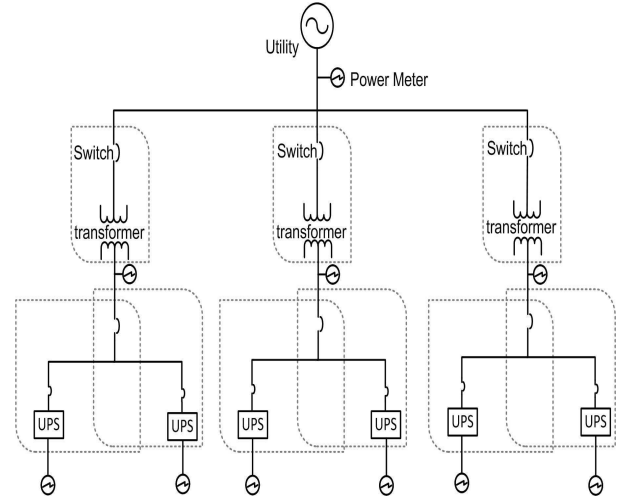In order to classify flow of current as data (integer values), we divide the time into small slices where in each slice we classify the flow of power on any link as integer values. For the sake of generality, we represent these classes from c_1 to c_n. The power meters measure power quality continuously and assign a power quality class to each time slice according to the power quality in the interval. So in every time slice, we know power quality values in terms of integers on links where the smart meters are installed. Our objective is to correctly estimate the power quality values on links where no smart meter is installed. Moreover, in order to simplify our model, we treat the power flow through each node as a channel (shown in Figure 2). The input and output of this channel at each node comprises n power quality classes. The probability that a power quality c_x will be received as c_y at the output of the channel at each device d is represented by the symbol $p^{(d)}_{c_y|c_x}$. For each device d (or a subnet consisting of the concatenation of several devices), we call the $n \times n$ matrix consisting of the probability values $p^{(d)}_{c_y|c_x}$ the *power*

(a) A subnet containing 2 devices.

(b) Subnet as a single power transition matrix.
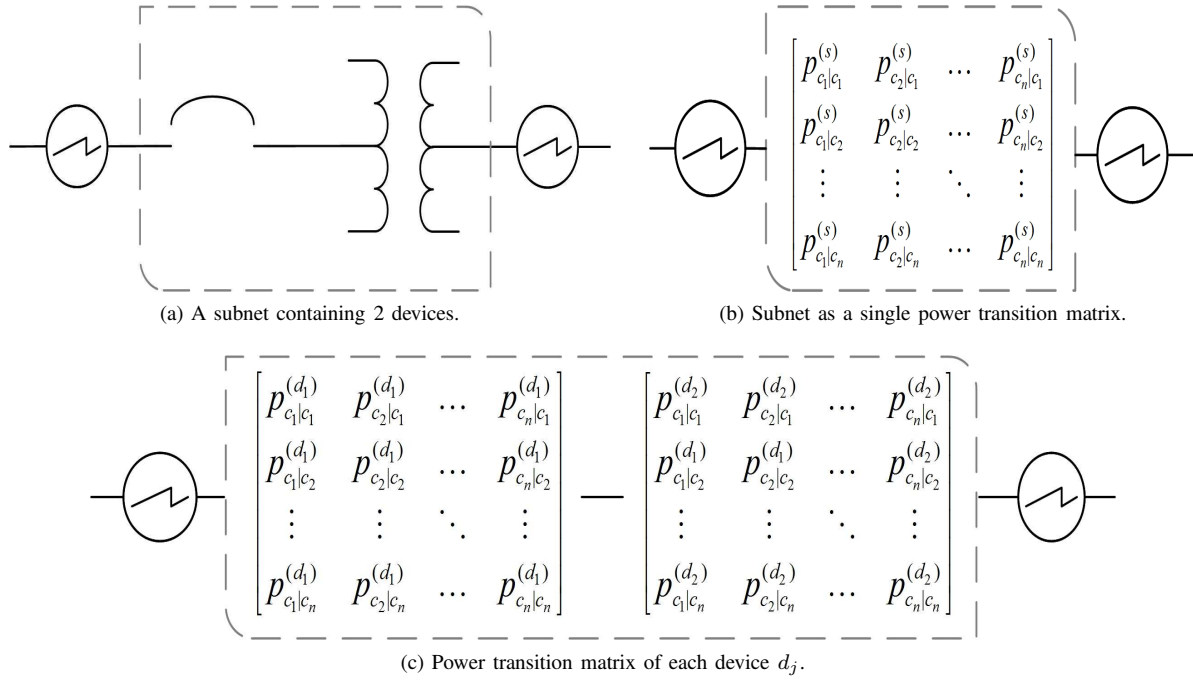
(c) Power transition matrix of each device d_j.

Fig. 4. Subnet power transition matrix as a product of power transition matrices of individual devices.

quality transition function[1], or simply *transition function*.

Figure 1 shows a view of a power grid where there are different types of electrical devices connected to each other via power links. The smart meters are also installed on selected links. Moreover, every type of device has a power quality function which may be unknown.

We want to estimate all the power quality functions based on the power quality values available on selected links where smart meters are installed. We divide the power grid network into small subnets where in every subnet we have two smart meters, i.e., one at the input as the first node and one at the output as the last node of the subnet as shown in Figure 3.

Now, we consider every subnet as a single node which we call a black-box. Figure 4 shows the view of a subnet as a single node. For every black-box, we know the power quality values at the input as well as at the output. Based on this known power quality information, we calculate the power quality transition function by using some sample readings from the smart meters attached to both sides of every black-box. We represent the calculated power quality function $f(s)$ of subnet s as a matrix as follows:

$$f(s) = \begin{bmatrix} p^{(s)}_{c_1|c_1} & p^{(s)}_{c_2|c_1} & \cdots & p^{(s)}_{c_n|c_1} \\ p^{(s)}_{c_1|c_2} & p^{(s)}_{c_2|c_2} & \cdots & p^{(s)}_{c_n|c_2} \\ \vdots & \vdots & \ddots & \vdots \\ p^{(s)}_{c_1|c_n} & p^{(s)}_{c_2|c_n} & \cdots & p^{(s)}_{c_n|c_n} \end{bmatrix}, \quad (1)$$

where $p^{(s)}_{c_y|c_x}$ is the probability that the input quality c_x is received as c_y at the output of the subnet s. Moreover, the

[1]For a device (or subnet) having multiple inputs/outputs, a power quality transition function is associated with each input/output pair.

value of $p^{(s)}_{c_y|c_x}$ can be easily calculated by measuring power quality values at the input and output smart meters of the subnet. Note that every row in the above matrix must sum to 1.

Figure 4(a) shows a subnet containing two devices. The power quality transition function (as a matrix) of the subnet is shown inside the black-box in Figure 4(b). Our objective is to correctly estimate the transition functions of every device d_j in the subnet. These unknown functions are shown as matrices in Figure 4(c).

In order to estimate the reliability of every device in the network, we need to estimate the power quality function $f(d_j)$ for each device d_j based on the quality function $f(s)$ of the subnet. It is clear that

$$f(s) = \prod_j f(d_j), \quad (2)$$

which implies that

$$\begin{bmatrix} p^{(s)}_{c_1|c_1} & \cdots & p^{(s)}_{c_n|c_1} \\ \vdots & \ddots & \vdots \\ p^{(s)}_{c_1|c_n} & \cdots & p^{(s)}_{c_n|c_n} \end{bmatrix} = \prod_j \begin{bmatrix} p^{(d_j)}_{c_1|c_1} & \cdots & p^{(d_j)}_{c_n|c_1} \\ \vdots & \ddots & \vdots \\ p^{(d_j)}_{c_1|c_n} & \cdots & p^{(d_j)}_{c_n|c_n} \end{bmatrix} \quad (3)$$

Our objective is to estimate $p^{(d_j)}_{c_y|c_x}$ (probability that power quality c_x will be mapped to power quality c_y at each device d_j). We use the data estimation techniques to solve Eq. (3).

IV. Power Quality Estimation using Entropy Maximization

Modeling the microgrids as data-driven networks helps us utilize the best available data estimation techniques to solve the power quality estimation problem. One of the most popular and widely-used data estimation techniques is EM algorithm. In a recent study [7], the authors successfully applied the EM algorithm to power quality estimation in microgrid. Although effective, the EM algorithm took a long time to converge. In this paper, we propose to use a MaxEnt approach to estimating the power quality of unmonitored segments of the grid. Before we formulate the power quality estimation problem as a MaxEnt problem, first we give a short description of EM and MaxEnt algorithms:

- **The Expectation Maximization (EM):** EM is a general approach to iterative computation of maximum-likelihood estimates when the observations can be viewed as incomplete data. Since each of the iteration of the algorithm consists of an expectation step followed by a maximization step, the algorithm is named as the EM algorithm. The successive iterations always increase the likelihood and the algorithm converges at a stationary point.

- **MaxEnt:** MaxEnt solves convex optimization problems of the form,

$$\textbf{maximize } g(\vec{x}) = -\sum_{i=1}^{n} x_i \log x_i$$

$$\text{subject to } \mathbf{A}\vec{x} \leq \mathbf{c}, \ \mathbf{B}\vec{x} = \mathbf{1},$$

where $\vec{x} \in \mathbb{R}^n$ is the optimization variable, $A \in \mathbb{R}^{m \times n}$, and $B \in \mathbb{R}^{m \times n}$ are problem parameters; and $\mathbf{1}$ is a vector with all 1's.

Now, we model the power quality estimation problem as the MaxEnt problem to accurately estimate the power quality transition functions $f(d_j)$ with acceptable efficiency. Obviously, there are multiple possible power quality functions $f(d_j)$ which are consistent with Eq. (3). We consider only those solutions which not only satisfy Eq. (3) but are consistent with other design constraints, for instance: 1) every row in the matrix $f(d_j)$ must sum to 1 (or at least very close to 1); and 2) $p_{c_y|c_x}^{(d_j)}$ must not be negative. Other possible constraints are discussed later in this section.

The basic idea of using MaxEnt here is that out of all possible quality functions (probability distributions) consistent with the design constraints, we choose the one with maximum uncertainty. Intuitively, the principle of MaxEnt implies that we should make use of all the information (design constraints) that is available and avoid making (biased) assumptions about information that is not available. Our objective function becomes,

$$\textbf{maximize } g = -\sum_{j, \, x, \, y} p_{c_y|c_x}^{(d_j)} \log p_{c_y|c_x}^{(d_j)} \tag{4}$$

subject to following constraints:

TABLE I
TRANSITION FUNCTIONS OF VARIOUS ELECTRICAL COMPONENTS [7]

		Output PQ				
		Class 1	Class 2	Class 3	Class 4	Class 5
Input PQ	Class 1	0.6	0.1	0.1	0.1	0.1
	Class 2	0.4	0.3	0.1	0.1	0.1
	Class 3	0.4	0.1	0.3	0.1	0.1
	Class 4	0.4	0.1	0.1	0.3	0.1
	Class 5	0.4	0.1	0.1	0.1	0.3

(a) Switch

		Output PQ				
		Class 1	Class 2	Class 3	Class 4	Class 5
Input PQ	Class 1	0.85	0	0	0	0.15
	Class 2	0.35	0.50	0.15	0	0
	Class 3	0.20	0.15	0.50	0.15	0
	Class 4	0.20	0	0.15	0.50	0.15
	Class 5	0.35	0	0	0.15	0.50

(b) Transformer

		Output PQ				
		Class 1	Class 2	Class 3	Class 4	Class 5
Input PQ	Class 1	1	0	0	0	0
	Class 2	0.1	0.9	0	0	0
	Class 3	0.1	0	0.9	0	0
	Class 4	0.1	0	0	0.9	0
	Class 5	0.1	0	0	0	0.9

(c) Bus

		Output PQ				
		Class 1	Class 2	Class 3	Class 4	Class 5
Input PQ	Class 1	1	0	0	0	0
	Class 2	0.8	0.2	0	0	0
	Class 3	0.8	0	0.2	0	0
	Class 4	0.8	0	0	0.2	0
	Class 5	0.8	0	0	0	0.2

(d) UPS

1) All components of the estimated functions $f(d_j)$ must not be negative, that is,

$$p_{c_y|c_x}^{(d_j)} \geq 0 \tag{5}$$

2) Every row of $f(d_j)$ sum to 1. To allow negligible rounding errors, we introduce a small rounding error factor (ε_1) to be tolerated, i.e.,

$$\left| 1 - \sum_y p_{c_y|c_x}^{(d_j)} \right| \leq \varepsilon_1, \quad \forall x \tag{6}$$

3) The estimated functions $f(d_j)$ satisfy the condition $\prod_j f(d_j) = f(s)$. One may relax the condition by a small approximation error factor (ε_2) to adjust the rounding errors. This adjustment is necessary when we are interested in a close approximation of $f(s)$. This condition becomes,

$$\left| f(s) - \prod_j f(d_j) \right| \leq \varepsilon_2. \tag{7}$$

Note that we slightly abuse the notation for simplicity. We use $-$ above to represent an element wise minus of two matrices and $|.|$ to change the values in the matrix to
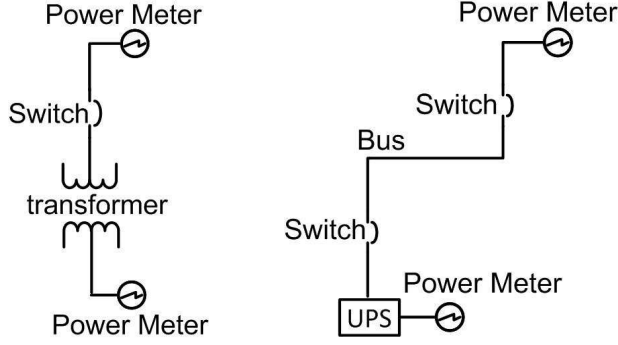
Fig. 5. Two different kinds of subnets. The one at the left side containing two devices (switch, transformer) and the one shown at the right side containing four devices (switch, bus, switch, and UPS).

		MaxEnt	EM		
			Iteration 5	Iteration 10	Iteration 15
Subnet size	2	2.16	40.3	117.7	402.6
	4	3.42	118.32	366.9	981.01

		MaxEnt	EM		
			Iteration 5	Iteration 10	Iteration 15
Subnet size	2	0.00020	0.00017	0.00013	0.00010
	4	0.00030	0.00025	0.00019	0.00014

their absolute values. In addition, the symbol $\leq$ means all values in the left matrix is no larger than ε_2.

4) Every component of the estimated function $f(d_j)$ must not vary from its corresponding component of the "true" transition function $\acute{f}(d_j)$ by a factor ε_3, i.e.,

$$\left| f(d_j) - \acute{f}(d_j) \right| \leq \varepsilon_3 \tag{8}$$

The meanings of the notations are the same as above.

Remark 1: In practice, we do not know $\acute{f}(d_j)$. To avoid this problem, we can *initially* set $\acute{f}(d_j)$ as the transition function of the same type of device, which may be obtained via its specification or historical data of monitored devices of the same type. As time goes, this initial setting of $\acute{f}(d_j)$ should be updated with the optimal estimation function, i.e., $\acute{f}(d_j)$ is set to equal $f(d_j)$, and the updated $\acute{f}(d_j)$ is used for the next round of estimation. Such iterative updates capture the transition function of the (unmonitored) device, which may change over time.

Remark 2: Someone may argue that our solution may be biased toward the last constraint (i.e., Eq. 8). By ignoring the last constraint, we get many solutions consistent to the first three constraints and our objective function chooses the one having maximum uncertainty. Here, if we do not use constraint 4 (Eq. 8), we will be ignoring some known information about the transition functions of the unmonitored devices which may result in compromising the accuracy of the estimated transition functions. Further, in some rare cases, the last constraint may not be feasible when an obtained transition functions ($f(d_j)$) is different by an amount greater than ε_3. In that case, we may ignore the last constraint or increase the value of ε_3.

Remark 3: It is worth noting that the objective function (4) is a non-Shannon measure, because $\sum_{j,\ x,\ y} p_{c_y|c_x}^{(d_j)} = \sum_j \sum_x \sum_y p_{c_y|c_x}^{(d_j)} > 1$. Nevertheless, by maximizing this non-Shannon measure, we obtain a good estimation of transition functions as demonstrated in our experimental evaluation. This is not by coincidence, since the objective function could be considered as the sum of several Shannon entropy functions (i.e., Given x and j, $\sum_y p_{c_y|c_x}^{(d_j)} \log p_{c_y|c_x}^{(d_j)}$ is a Shannon entropy measure). In our case, we treat each entropy function with the same weight. If we have more knowledge regarding the distribution of input power quality events, assigning different weights to different entropy functions may lead to better estimation. We leave this exploration as future work.

Remark 4: This paper focuses on a centralized processing and as such we assume that there exists an underlying communication system to support data collection to a central point. The implementation of this communication system is beyond the scope of this paper. Further, the real-time issues of the monitoring and communication systems such as the data sampling and communication delay are ignored in the paper.

V. PERFORMANCE EVALUATION

We implement the proposed objective function and simulate a power microgrid using MATLAB. The view of the simulated network is shown in Figure 1 where only several network segments are monitored using smart meters. Our objective is to estimate the power quality values on network segment where no smart meter is installed. We use the same transition functions as in [7] as the ground truth. These power quality transition functions of various electrical components (switch, bus, ups, and transformer etc) in the smart grid are shown in Table I.

We use non-linear constrained optimization algorithm named Sequential Quadratic Programming (SQP) to estimate the unknown power quality functions. Inputs to the algorithm are the known power quality functions for each subnet $f(s)$, error tolerance factors $\varepsilon_1 = \varepsilon_2 = 0.01$, and $\varepsilon_3 = 0.05$. The error tolerance is small enough for practical applications. The power quality matrix of each subnet is the product of the power transition functions (shown in Table I) of the devices in that subnet. The simulated microgrid consists of subnets of two different sizes; one contain two devices (switch and a transformer) while the other containing four devices (switch, bus, switch, and UPS). Both of these subnets are shown in Figure 5. The power transition matrices of each subnet are calculated by multiplying the ground truth matrices of devices used in the corresponding subnet. Further, the simulation was performed on a desktop computer having Intel Dual-Core-i7 3.4 GHz processor with 4 GB physical memory.

Table II shows the convergence time comparison of both the

EM and MaxEnt based solutions to power quality estimation. It can be seen that the convergence time of MaxEnt is much faster as compared to that of EM algorithm. We measure the estimation accuracy using Mean Squared Error (MSE), which is a statistical measure quantifying the difference between values implied by an estimator and the true values of the quantity being estimated. The results are shown in Table III. From the results, we can see that both the methods give very close estimations of the power quality transition functions, i.e., the difference between the estimated and corresponding ground truth functions is negligible.

VI. CONCLUSIONS AND FUTURE WORK

Reliability of power grid networks is critically important in today's life. Smart meters play an important role in estimating reliability of power networks but they are expensive devices and it is financially infeasible to install them on every link between devices in the power grid network. We propose a MaxEnt based model that accurately estimates power quality transition functions on unmonitored network segments. The experimental results show that 1) our MaxEnt based solution is much faster than the existing EM based solution to power quality estimation; and 2) the proposed solution accurately estimates the power transition functions.

Finally, the MaxEnt based model opens a new scope of methods to quantitatively measure and solve the reliability problems in smart grid. As our future work, we will study the optimal measurement strategy using the MaxEnt approach.

REFERENCES

[1] Fluke Corporation, "Fluke Power Quality Analyzers," http://www.fluke.com/fluke/usen/products/categorypqttop.htm, [Online: accessed Aug. 2013].

[2] Schneider Electric, "PowerLogic Energy and Power Quality Meters," http://www.schneider-electric.com, [Online: accessed Aug. 2013].

[3] E. T. Jaynes, "Information theory and statistical mechanics," *The Physical Review*, vol. 106, no. 4, pp. 620–630, 1957.

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, B*, vol. 39, 1977.

[5] IEEE, "IEEE recommended practice for the design of reliable industrial and commercial power systems," *IEEE Std 493-2007 (Revision of IEEE Std 493-1997)*, pp. 1–383, 2007.

[6] I. T. I. C. (ITI), "ITI (CBEMA) Curve Application Note," Information Technology Industry Council (ITI), Tech. Rep., October 2007.

[7] C. Gamroth, K. Wu, and D. Marinakis, "A smart meter based approach to power reliability index for enterprise-level power grid," in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. IEEE, 2012, pp. 534–539.

[8] IEEE, "IEEE recommended practice for monitoring electric power quality," *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, pp. c1–81, June 2009.