IBM Research Report

Iterative Bayesian Demand Response Estimation and Sample-Based Optimization for Real-Time Pricing

Pu Huang
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

IBM

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Iterative Bayesian Demand Response Estimation and Sample-Based Optimization for Real-Time Pricing

Pu Huang, *Member, IEEE*

Abstract—Dynamic pricing will introduce significant load fluctuation in the wholesale markets than we see nowadays as customers respond to price signals in real-time. In this paper, we propose an iterative estimation-optimization approach for real-time wholesale market pricing in response to stochastic load fluctuations. For additive and multiplicative demand response models, we apply a Bayesian approach to continuously update the models. The estimated demand response models are then fed into an expectation-base or a risk-averse optimal flow problem to find bus phase angles, real power injections, and bus prices simultaneously. We also present preliminary numerical results using simulated data.

I. INTRODUCTION

Electricity markets in U.S. have a two-layered structure. Prices in many wholesale markets are set in a competitive way (e.g., PJM, ISO New England, and Midwest ISO adopted the Locational Marginal Pricing (LMP) approach), but for retail markets, the rates in general are still fixed. Rapid deployment of smart meters in recent years paved the way for dynamic pricing in retail markets. Dynamic pricing has been proposed as an effective mechanism to align customer energy usage decisions with cost responsibility. Dozen of pilot programs have been carried out across the global to experiment dynamic pricing, and significant benefits have been recorded [1], [2]. One side effect of dynamic pricing is that it will introduce significant load fluctuation in the wholesale markets than we see nowadays as customers respond to dynamic price signals in real-time.

In this paper, we propose an iterative estimation-optimization approach to adjust wholesale market prices in real-time in response to stochastic load fluctuations. To link demand and price, we assume a stochastic demand response model. In each time period, Bayesian updating is applied to update the parameters of the demand response model based on the realized price and demand data collected so far. Then the updated model is plugged into an extended Optimal Power Flow (OPF) formulation to find the optimize price for the next period, and the process continues.

Load forecasting has always been important issue in the power industry. A fairly substantial literature exists on this subject [3], [4], [5], [6], [7]. However, estimating the impact of price on load has not been of much interest until recently. This is, of course, understandable as price was fixed and not allowed to change before market deregulation. A few

dynamic pricing pilots carried out across the nation provided quality data to estimate price impacts (refer to [8], [1] and the references therein). However, most of these studies were poster experiment analysis and thus have been focused on descriptive statistics like demand elasticity, curtailed peak demand, etc., not on real-time demand response models. For our purpose, we investigate two commonly use demand response models: additive and multiplicative models. Both are simple and well known models that prescript linear (log-linear) relations between demand and price. We adopt a Bayesian approach to update model parameters continuously as new price and demand data become available. To find the optimal price, we extend the DC-OPF formulation to maximize the overall social benefit, while at the same time keep all the security constraints in the original formulation. Thus we simultaneously determine typical phase angles, real power injections, as well as prices at individual buses.

Since we can never forecast loads exactly, our optimization model needs to handle load uncertainty. We investigate two approaches: the traditional expectation-based approach, and a risk-averse approach. We propose to sample the posterior distributions after Bayesian updating to estimate the values of input quantities to the optimization model, and then feed the estimates into the subsequent optimization problem. Our treatment of load uncertainty differentiates itself from the Probabilistic OPF (POPF) approach [9], [10], [11], [12]. POPF seeks to find (or for that matter, estimate) the probability distribution functions (pdf) of the quantities of interest (phase angles, bus voltages, LMPs, etc.) for a given set of load distributions. We don't estimate the pdf of the outputs from our optimization models. Instead, we estimate the pdf of the inputs using the Bayesian approach, set up a risk measure to gauge the risk of load uncertainty, and then balance the risk with the cost via optimization. The specific risk measure we use in this paper is Conditional Value-at-Risk (CVaR). We adopt it because of the intuitive meaning and nice computational proprieties [15] of this risk measure, though other measures can be applied as well.

To summarize, the proposed estimation-optimization procedure consists of the following steps:

- Estimate parameters of the demand response model based on the latest price and load data available.
- Generate random samples to estimate the values of inputs to the optimization model.
- Feed the estimates to the optimization problem and solve it to find phase angles, real power injections, and prices at individual buses.
- Send the price signals out, collect the realized load from

each bus, and then loop back the first step.

The rest of the paper is organized as follows. Section II describes the Bayesian updating procedures for additive and multiplicative demand response models. In Section III, we elaborate two optimization models: expectation-based and risk-averse models, and their respective solution algorithms. Section IV reports a set of preliminary numerical experiments we carried to test our approach. Section V concludes the paper.

II. Bayesian demand response model updating

We consider two widely used demand response models: multiplicative and additive models: They can be written, respectively, as the follows

$$
\begin{aligned}
d &= \exp(-\alpha p + a + \delta) \\
d &= -\beta p + b + \eta
\end{aligned}
$$

In an on-line setting, they take the following form:

$$
\begin{aligned}
\log(d_t) &= -\alpha_t p_t + a_t + \delta, & (1) \\
d_t &= -\beta_t p_t + b_t + \eta, & (2)
\end{aligned}
$$

where t indexes time, α_t and β_t are *estimated* demand response parameters in their respective model, δ and η are error terms that don't vary with time t, and a_t and b_t are *known* quantities that influence demand (e.g. demand in period $t-2$, weather information, demand in the same time period day/week/month ahead, etc.). a_t and b_t are function of t, but for our modeling purpose, we do not treat them as random variables as we want to single out the impacts of *unknown* parameters α_t, β_t, δ, and η.

Prices in real-time wholesale markets typically are updated in a fixed frequency. For example, prices in the PJM market are posted in every 5 minutes [13]. In every periods t, the following events occur in sequence:

- Actual load d_{t-1} in the previous period $t-1$ is recorded,
- The demand response model is updated based on actual load and price information,
- Price p_t for this period is posted,
- Continue to the next period.

Both demand response models (1) and (2) are special cases of the following general linear model

$$
y = kx + c + \epsilon, \tag{3}
$$

where x and y are the input and output of the model, k and c are parameters, and $\epsilon \sim Normal(0, \sigma^2) = Normal(0, h^{-1})$ is an error term. For multiplicative model (1),

$$
y := \log(d_t), x := p_t, k := -\alpha_t, c := a_t, \epsilon := \delta;
$$

and for additive model (2),

$$
y := d_t, x := p_t, k := -\beta_t, c := b_t, \epsilon := \eta.
$$

Let $\mathcal{D} = (y_n - c_n, x_n), n = 1, ..., N$ denote N observations of (y, x), assign the conjugate prior to k and h, then the following Bayesian updating rules are standard [14]:

- Prior of h and k is specified as

$$
\begin{aligned}
h &\sim Gamma(\underline{\nu}/2, \frac{2}{\underline{s}^2 \underline{\nu}}) & (4) \\
k|h &\sim Normal(\underline{k}, h^{-1}\underline{V}) & (5)
\end{aligned}
$$

- and after Bayesian updating, the posterior is

$$
\begin{aligned}
h|\mathcal{D} &\sim Gamma(\bar{\nu}/2, \frac{2}{\bar{s}^2 \bar{\nu}}) & (6) \\
k|h, \mathcal{D} &\sim Normal(\bar{k}, h^{-1}\bar{V}) & (7)
\end{aligned}
$$

See Appendix A for more details on how parameters $\bar{\nu}, \bar{s}, \bar{k}$ and $\bar{V}$ of the posterior can be computed.

We carry out the above Bayesian update in every period t. Since we get only one sample point (d_t, p_t) in every period, to apply the Bayesian approach, we backtrack another $N-1$ periods to get N points in total.

III. Optimization Formulations and Solution Algorithms

Our optimization formulation is based on the DC-OPF model. Define the following notation:

- $\mathcal{I}$: set of buses/generators, assume one generator per bus,
- $\mathbf{g}$: real power injection vector,
- $\mathbf{d}$: bus real power load vector,
- $\mathbf{B}$: bus admittance matrix (imaginary part),
- $\boldsymbol{\theta}$: bus voltage angle vector,
- $\mathcal{L}$: set of transmission lines, indexed by ij,
- l_{ij}: power flow limit of transmission line between buses i and j,
- x_{ij}: reactance of line ij.
- θ_i^{min} and θ_i^{max}: lower and upper bounds of bus i voltage angle,
- g_i^{min} and g_i^{max}: minimum and maximum power output of generator i,
- $c_i(\cdot)$ cost function of generator i.

The basic DC-OPF model can be written as,

$$
\begin{aligned}
\min_{g_i, \theta_i} &\sum_{i \in \mathcal{I}} c_i(g_i) & (8a) \\
&\mathbf{B}\boldsymbol{\theta} \leq \mathbf{g} - \mathbf{d} & (8b) \\
&\left| \frac{\theta_i - \theta_j}{x_{ij}} \right| \leq l_{ij} & \forall ij \in \mathcal{L} & (8c) \\
&\theta_i^{min} \leq \theta_i \leq \theta_i^{max} & \forall i \in \mathcal{I} & (8d) \\
&g_i^{min} \leq g_i \leq g_i^{max} & \forall i \in \mathcal{I}. & (8e)
\end{aligned}
$$

With demand response, we would like to utilize dynamic prices as a "control signal" to modulate the demand. To this end, we first write down the (estimated) response function that links demand d_t and price p_t in each period t, and then plug this function into an optimization problem (see problem (10a) below) to find the optimal price. For the multiplicative model, the demand response function takes the form (1), and for the additive model, it takes the form (2). Let

$$
d = f_{gen}(p|\omega) \tag{9}
$$

denote a general demand response model with a vector of stochastic parameters ω (represent α_t, β_t, δ, and η), we would like to solve the following extension of the basic DC-OPF

model to determine the optimal price:

$$\min_{g_i,\theta_i,p_i} \sum_{i\in\mathcal{I}} c_i(g_i) - \sum_{i\in\mathcal{M}} p_i d_i \tag{10a}$$

$$\mathbf{B}_{(i,\bullet)}\boldsymbol{\theta} \leq g_i \qquad i \in \mathcal{I} - \mathcal{M} \tag{10b}$$

$$\mathbf{B}_{(i,\bullet)}\boldsymbol{\theta} \leq g_i - d_i \quad i \in \mathcal{M} \tag{10c}$$

$$\left|\frac{\theta_i - \theta_j}{x_{ij}}\right| \leq l_{ij} \qquad \forall ij \in \mathcal{L} \tag{10d}$$

$$\theta_i^{min} \leq \theta_i \leq \theta_i^{max} \qquad \forall i \in \mathcal{I} \tag{10e}$$

$$g_i^{min} \leq g_i \leq g_i^{max} \qquad \forall i \in \mathcal{I} \tag{10f}$$

$$d_i = f_{gen}(p_i|\omega_i) \qquad \forall i \in \mathcal{M} \tag{10g}$$

where $\mathcal{M} \subset \mathcal{I}$ is a subset of buses that have loads, and $\mathbf{B}_{(i,\bullet)}$ is the i-th row of $\mathbf{B}$.

In the above optimization model, we optimize the social welfare, and introduce new decision variables p_i. Note that load d_i now becomes a random function of p_i. Because of the random natural of d_i, problem (10a) is not well defined. For a given control p_i, we don't know exactly what the outcome d_i would be (though we do know the pdf of d_i).

A. Expectation Optimization

One traditional way to handle random variable d_i is to replace it with its mean. Doing so leads to the following problem

$$\min_{g_i,\theta_i,p_i} \sum_{i\in\mathcal{I}} c_i(g_i) - \sum_{i\in\mathcal{M}} p_i d_i \tag{11a}$$

$$(10b),(10c), (10d), (10e), (10f) \tag{11b}$$

$$d_i = E[f_{gen}(p_i|\omega_i)] \qquad \forall i \in \mathcal{M}, \tag{11c}$$

More specifically, for multiplicative model (1)

$$d_{i,t} = E[\exp(-\alpha_{i,t}p_{i,t} + \delta_i)]\exp(a_{i,t}) \tag{12}$$

and for additive model (2),

$$d_{i,t} = E[-\beta_{i,t}]p_{i,t} + b_{i,t} + E[\eta_i]$$
$$= E[-\beta_{i,t}]p_{i,t} + b_{i,t} \tag{13}$$

Here we use double index (i,t) to identify demand $d_{i,t}$ at bus $i \in \mathcal{M}$ in time period t. Since we solve an optimization like (11a) in every period, index t is omitted to keep the notation simple. Error term η_i has zero mean, and hence is ignored in (13). However, we have to keep δ_i in (12) as it is an exponent.

B. Risk-Averse Optimization

One issue with formulation (11a) is that it assumes *average* demand response (as evident by the use of expectation operator $E[\cdot]$). We may also want to stay "on the safe side" by, say, considering a "worse-than-average" demand response scenario. This intuition of risk-averse can be modeled precisely in a probabilistic framework using a risk measure call Conditional Value-at-Risk (CVaR).

For a given random variable X, the conditional value-at-risk of X, $\rho_{1-\gamma}[X]$, parameterized by $1 - \gamma$, maps X to a real number. Specifically, $\rho_{1-\gamma}[X]$ is the expected value of X given that X is greater than its $(1 - \gamma)$-quantile, i.e.,

$$\rho_{1-\gamma}[X] = E[X|X \geq q_{1-\gamma}(X)],$$

where $q_{1-\gamma}[X]$ represents the $(1-\gamma)$-quantile of X. By definition, $\rho_{1-\gamma}[X]$ measures the "average" of the worst possible realizations of X (X here represents losses, so the larger the worse), where the γ controls what values are considered as the "worst realizations". For example, $\gamma = 5\%$ means that top 5% of all the values X may take are counted as the "worst realizations".

Replace expectation $E[\cdot]$ in formulation (11a) by $\rho_{1-\gamma}[\cdot]$, we get the following risk-aware optimization formulation

$$\min_{g_i,\theta_i,p_i} \sum_{i\in\mathcal{I}} c_i(g_i) - \sum_{i\in\mathcal{M}} p_i d_i \tag{14a}$$

$$(10b),(10c), (10d), (10e), (10f) \tag{14b}$$

$$d_i = \rho_{1-\gamma_i}[f_{gen}(p_i|\omega_i)] \qquad \forall i \in \mathcal{M}. \tag{14c}$$

Again, specific functional forms of the multiplicative and additive models are, respectively

$$d_{i,t} = \rho_{1-\gamma_i}[\exp(-\alpha_{i,t}p_{i,t} + \delta_i)]\exp(a_{i,t}) \tag{15}$$

$$d_{i,t} = \rho_{1-\gamma_i}[-\beta_{i,t}p_{i,t} + \eta_i] + b_{i,t} \tag{16}$$

Note that since

$$\rho_{1-\gamma}[\lambda X] = \lambda\rho_{1-\gamma}[X], \quad \forall \lambda \geq 0,$$

we can take $\exp(a_{i,t})$ out of the risk function in (15). In general, we cannot distribute operator $\rho_{1-\gamma}[\cdot]$ to individual random variables $\beta_{i,t}$ and η_i in (16), as it is not linear as $E[\cdot]$. Indeed, it is a convex operator, i.e., for two random variables X and Y, we have

$$\rho_{1-\gamma}[X + Y] \leq \rho_{1-\gamma}[X] + \rho_{1-\gamma}[Y] \tag{17}$$

Refer to [15] for details about the properties of the CVaR risk measure.

C. Solution Algorithms

We elaborate our solution algorithms for the expectation-based problem (11a) and for the risk-aware optimization (14a)

1) Algorithm for Expectation-Based Formulation: For the additive model, plugging (13) into problem (11a) leads to a quadratic optimization problem with linear constraints. It is straightforward to solve it using any off-the-shelf quadratic programming solvers. $E[-\beta_{i,t}]$ in equation (13) is available in every period t after Bayesian updating. It equals to $\bar{k}$ as in the general linear representation (3). In this paper, we generate a set samples for $\beta_{i,t}$ at each bus $i \in \mathcal{M}$ and then feed the sample averages as the inputs into optimization problem (11a). We could use $\bar{k}$ directly, but a sample-based approach is more generally applicable, as it is always possible to generate samples for any quantities of interest in an Bayesian framework. For all the algorithms we are going to discuss in the following, we take empirical sample estimates as inputs to the corresponding optimization models.

Problem (11a) with multiplicative demand response model (12) contains a set of exponential functions. One way of handling them is to use nonlinear solvers. Instead of doing that, we adopt an approximation approach and use readily available

quadratic solvers for solution. Ignore bus index i and rewrite (12) as

$$
\begin{aligned}
d_t &= E[\exp(-\alpha_t p_t + \delta)] \exp(a_t) \\
&\approx E[1 - \alpha_t p_t + \delta] \exp(a_t) \\
&= (1 + E[-\alpha_t] p_t) \exp(a_t) \quad (18)
\end{aligned}
$$

Here we use the first-order Taylor expansion to approximate the exponential function around zero. Notice that (18) is linear and thus plugging it to problem (11a) again leads to a quadratic program with all linear constraints. Again, $E[-\alpha_t]$ is estimated using samples generated from the posterior.

2) Algorithm for Risk-Averse Formulation: Ignore bus index i and re-write (16) as

$$
\begin{aligned}
d_t &= \rho_{1-\gamma}[-\beta_t p_t + \eta] + b_t \\
&\leq \rho_{1-\gamma}[-\beta_t] p_t + \rho_{1-\gamma}[\eta] + b_t \quad (19)
\end{aligned}
$$

The above inequality is a consequence of (17). In this paper, we assume d_t equals the r.h.s of (19) to price conservatively (i.e. assume demand is less responsive than the average case as in the expectation-based model). We generate random samples from the posterior to estimate $\rho_{1-\gamma}[-\beta_t]$ and $\rho_{1-\gamma}[\eta]$. To this end, let pair $(\beta_{t,s}, \eta_s)$, $s = 1, ..., S$, denote an i.i.d. sample with size S drawn from the posterior of (β_t, η). Then $\rho_{1-\gamma}[-\beta_t]$ and $\rho_{1-\gamma}[\eta]$ can estimate using the top γS samples from their respective margins. More precisely, sort η_s in ascending order, let $\eta_{(s)}$ denote the sorted η_s, then $\rho_{1-\gamma}[\eta]$ can be estimated as

$$
\frac{1}{S - \lceil S(1-\gamma) \rceil + 1} \sum_{s=\lceil S(1-\gamma) \rceil}^{S} \eta_{(s)}
$$

where $\lceil S(1-\gamma) \rceil$ denotes the smallest integer that is greater than $S(1-\gamma)$. $\eta_{\lceil S(1-\gamma) \rceil}$ is the estimated $(1-\gamma)$-quantile of η, and the above estimate literally is the conditional average of samples given that they are greater than the (estimated) $(1-\gamma)$-quantile of η, which is precisely the definition of CVaR. $\rho_{1-\gamma}[-\beta_t]$ can be estimated in a similar way.

Algorithms exist to handle the original CVaR constraint without approximating it using its upper bound, refer to [15], [16] for details.

For multiplicative demand response model (15), apply the same approximation method as described in (18), we have

$$
\begin{aligned}
d_t &= \rho_{1-\gamma}[\exp(-\alpha_t p_t + \delta)] \exp(a_t) \\
&\approx \rho_{1-\gamma}[1 - \alpha_t p_t + \delta] \exp(a_t) \\
&\leq (1 + \rho_{1-\gamma}[-\alpha_t] p_t + \rho_{1-\gamma}[\delta]) \exp(a_t) \quad (20)
\end{aligned}
$$

Again, we generate random samples to estimate $\rho_{1-\gamma}[-\alpha_t]$ and $\rho_{1-\gamma}[\delta]$.

IV. NUMERICAL EXPERIMENTS

A. Bayesian Update

We first show a simple example of applying the Bayesian rules to estimate a general linear as shown in (3). Set a non-informative prior [1] as

$$
\underline{\nu} = 1, \underline{s}^2 = 1000, \underline{k} = 0, \underline{V} = 1000,
$$

[1] Note that $\underline{s}^2 = 1000 \gg 4^2$, small $\underline{\nu}$ and large $\underline{V}$ values.
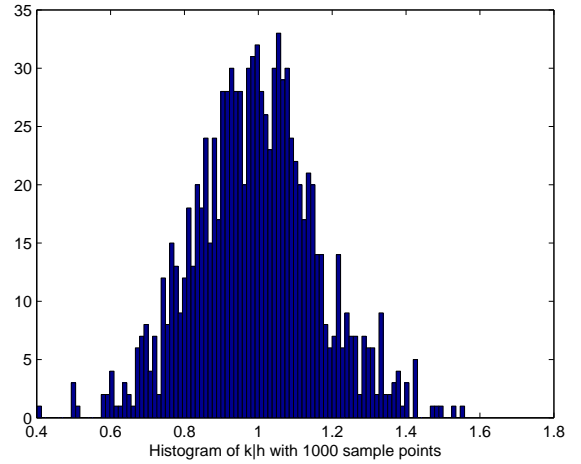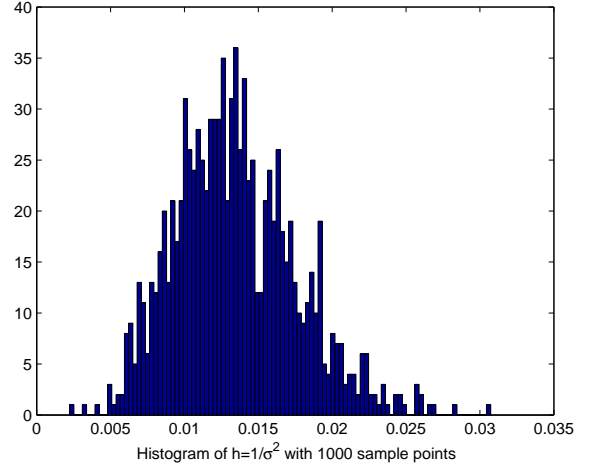

Fig. 1. Histogam of marginal posterior distributions of h and k with 1000 sample points.

$$
k = 1, \epsilon \sim Normal(0, \sigma^2) = Normal(0, 4^2),
$$

Generate $N = 20$ sample points, and the posterior values of the parameters are

$$
\bar{\nu} = 21, \bar{s}^2 = 74.9928, \bar{k} = 0.9936, \bar{V} = 3.4843e - 004.
$$

We further compute the following quantities that are typically of interest,

$$
\begin{aligned}
E[k] &= \bar{k} = 0.9936 \\
Std[k] &= \sqrt{\bar{s}^2 \bar{V}} = 0.1616 \\
E[\sigma] &= \sqrt{\bar{s}^2} = 8.6598
\end{aligned}
$$

We use the posterior parameters to generate 1000 samples of (k, σ^2). Figure 1 shows the histograms of the margins of (k, σ^2). Use these samples, we estimate

$$
E[k] = 0.9936, Std[k] = 0.1675, E[\sigma] = 8.6265
$$

These sample-based estimates are very close to the exact values. This assures us that the sample-based optimization approach we adopt will generate reasonable accurate solutions.

B. Additive Demand Response Model with Expectation Optimization

We extended Matpower [17] to solve problems (11a) and (14a). For the following numerical experiments, we take the IEEE 9-bus system (called "case9") included in the Matpower package as the basis for our tests.

There are 3 buses in case9 that have loads, and the loads are $90, 100$ and 125 respectively. Denote these buses as "load bus i", $i = 1, 2, 3$, respectively. Let pd_i denote the original deterministic load. We set β_i equal to $0.01, 0.02$ and 0.03, $a_i = pd_i$, and $\eta_i \sim Normal(0, (0.05)^2)$. The parameters of the initial prior are

$$\underline{\nu} = 1, \underline{s}^2 = 1000, \underline{\beta} = -10, \underline{V} = 1000$$

for each $i = 1, 2, 3$. We run our estimation-optimization procedure 100 iterations (i.e, t=1,...,100). During each iteration, the Bayesian procedure backtracks $N = 30$ periods to update the posterior. We set $N = 30$ artificial data points to jump-start the Bayesian procedure. The (y, x) values of these 30 points are set as 1 (so they carry no information). We generate $S = 1000$ sample points for each bus i in each period to estimate $E[-\beta_{i,t}]$. It may happen that the empirical sample estimate of $E[-\beta_{i,t}]$ is positive. This typically occurs during early iterations when the Bayesian posteriors still carry a large estimation error. If this happens, we reset the empirical estimate to -10, its prior value.

Figure 2 shows the dynamics of price and demand as functions of t for each load bus. In Figure 2(a), dashed flat lines represent the true optimal prices if the demand response function at each bus is known, and the solid lines show the evolution of price during the estimation-optimization process. It is evident that the less responsive is the demand (corresponding to smaller β_i values), the higher the price. Figure 2(b) shows the realized random demands (dashed lines) and the estimated mean demands based on samples generated from the posteriors.

## C. Multiplicative Demand Response Model with Expectation Optimization

We set all $\alpha_i = 0.005 * pd_i$ for all i, $a_i = \log(pd_i)$, and $\delta_i \sim Normal(0, (0.05 * pd_i)^2)$. All the other settings are the same as in the proceeding additive case. The evolution of price and demand are shown in Figure 3

D. Additive Demand Response Model with Risk-Averse Optimization

We set $\gamma_i = 80\%$ for all i (i.e., we average top 80% of the samples, the result is a value greater than the mean). All the other parameters are the same as in the case of additive model with expectation optimization. It turns out that the estimation errors of β and η slow down convergence of the risk-averse optimization algorithm. To better estimate the CVaR values, we increase sample size S to 5000. We found that during early iterations, the CVaR estimates are pretty bad because of the presence large posterior errors. We took two extra steps to handle such errors: 1) Use the estimated expected values of $-\beta$
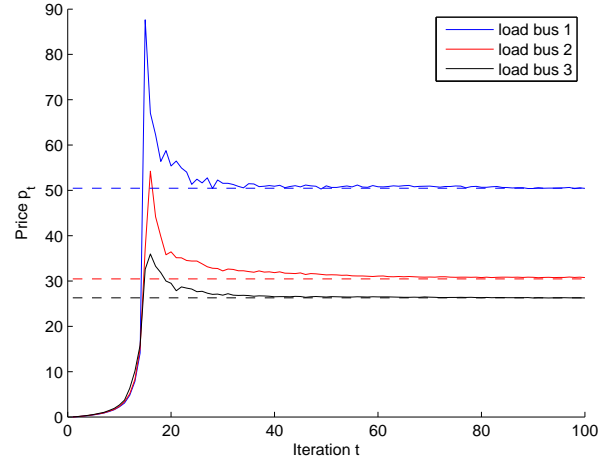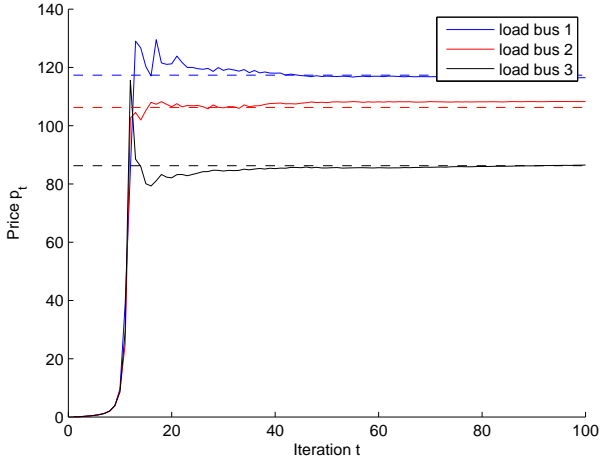

(a) Price
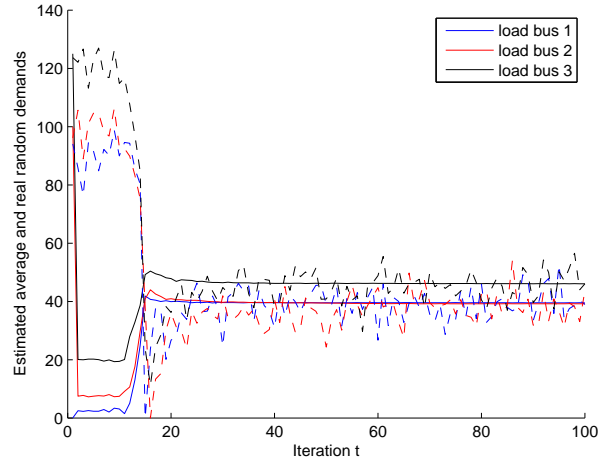
(b) Demand

Fig. 2. Price and demand dynamics: additive model with expectation optimization.

and η in the first 50 iterations. 2) In the following iterations, if the estimate of $\rho_{1-\gamma}[-\beta]$ is greater than 0, set it to -10; if the estimate of $\rho_{1-\gamma}[\eta]$ is less than 0, set it to 0. To make sure that prices converge, we increase the iteration number to 300. Figure 4 shows the price and demand dynamics. One observation is that compared with the expectation-based approach, the prices obtained by the risk-averse optimization are higher, which is consistent with our intuition that in our framework, less responsive demand is charged at a higher price. Figure 4(b) presents the mean demand estimation (i.e., $E[d_i]$, not $\rho_{1-\gamma}[d_i]$, for easy comparison, as the latter is always greater than the former) vs. realized demand.

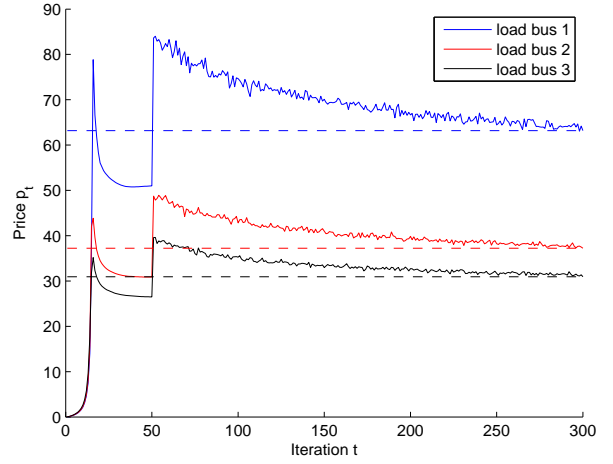## E. Multiplicative Demand Response Model with Risk-Averse Optimization

The settings are the same as in the expectation optimization case. We employ the same extra steps as in the proceeding case to handle large estimation errors of CVaR. For this experiment, we increase the iteration number to 1000 till the prices set
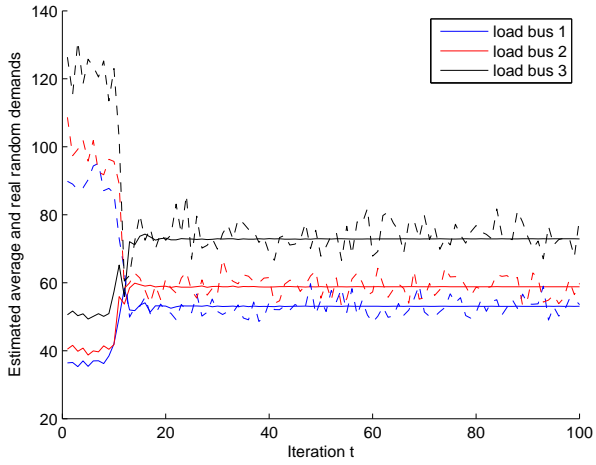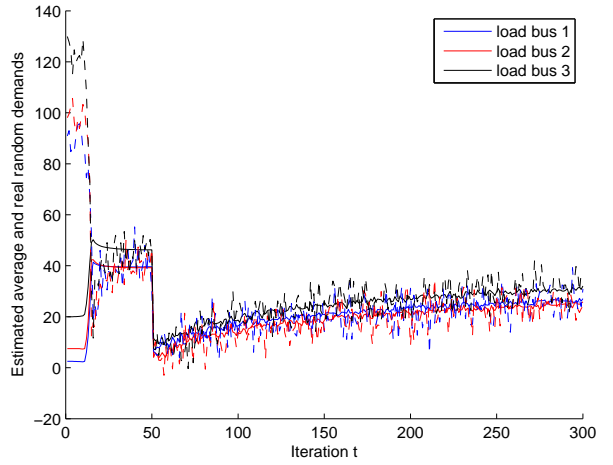
(a) Price

(b) Demand

Fig. 3. Price and demand dynamics: multiplicative model with expectation optimization.

(a) Price

(b) Demand

Fig. 4. Price and demand dynamics: additive model with risk-averse optimization.

down. The price and demand dynamics are shown in Figure 5.

V. CONCLUSION

We have presented an iterative approach to continuously estimate demand responses and optimize overall social welfare for wholesale market real-time pricing. One potential extension to this approach is to identify more realistic, yet computationally easy to handle demand response functions for large problem instance tests. In this paper, we solve a static optimization problem in each period, thus we ignore possible inter-temporal effects of demand response (e.g., one may curtail load in certain periods but still keep the total consumption unchanged). Another extension thus is to investigate such a dynamic model that explicitly counts for temporal demand responses.

APPENDIX A
COMPUTE PARAMETERS OF THE POSTERIOR

Given the prior distribution of (k, h) as specified in (4) and (5), and a set of sample points $\mathcal{D} = (y_n - c_n, x_n), n = 1, ..., N$, denote

$$\hat{k} = \frac{\sum_{n=1}^{N} x_n (y_n - c_n)}{\sum_{n=1}^{N} x_n^2}$$

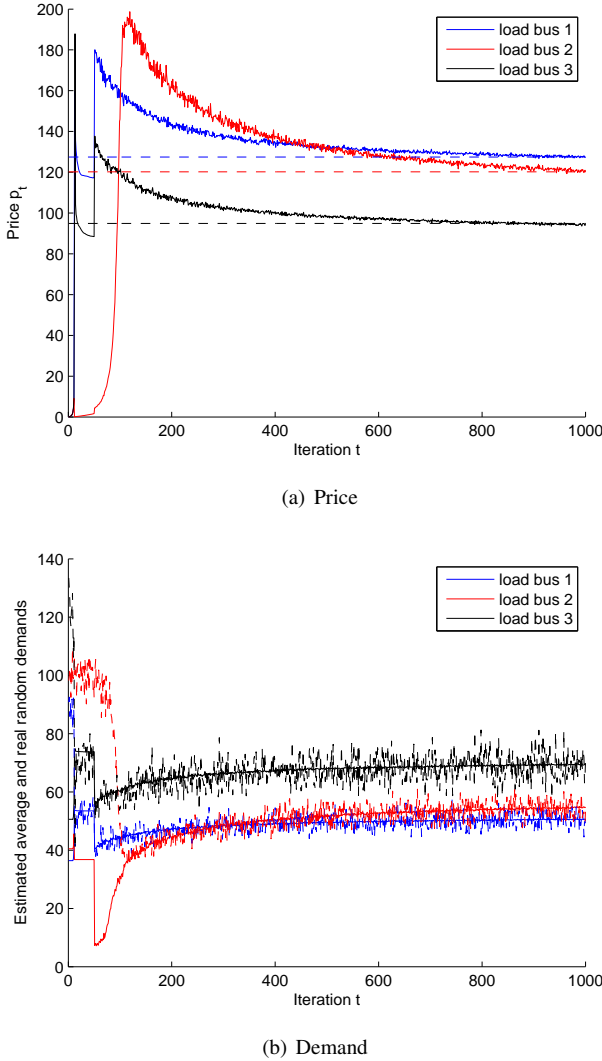$$s^2 = \frac{\sum_{n=1}^{N} (y_n - c_n - \hat{k} x_n)^2}{N - 1}$$

REFERENCES

[1] A. Faruqui and S. Sergici, "Household response to dynamic pricing of electricity: A survey of the experimental evidence," 2009, working Paper, Brattle Group.
[2] A. Faruqui, "The ethics of dynamic pricing," *The Electricity Journal*, vol. 23, no. 6, pp. 13 – 27, 2010.
[3] M. T. Hagan and S. M. Behr, "The time series approach to short term load forecasting," *Power Systems, IEEE Transactions on*, vol. 2, no. 3, pp. 785 –791, 1987.
[4] G. Gross and F. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558 – 1573, Dec. 1987.
[5] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *Power Systems, IEEE Transactions on*, vol. 4, no. 4, pp. 1484 –1491, Nov. 1989.
[6] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *Power Systems, IEEE Transactions on*, vol. 21, no. 1, pp. 392 – 401, feb. 2006.
[7] N. Amjady, "Short-term bus load forecasting of power systems by a new hybrid method," *Power Systems, IEEE Transactions on*, vol. 22, no. 1, pp. 333 –341, feb. 2007.
[8] H. Allcott, "Rethinking real-time electricity pricing," 2010, working Paper, Department of Economics, Massachusetts Institute of Technology.
[9] M. Madrigal, K. Ponnambalam, and V. Quintana, "Probabilistic optimal power flow," in *Electrical and Computer Engineering, 1998. IEEE Canadian Conference on*, vol. 1, May 1998, pp. 385 –388 vol.1.
[10] G. Verbic and C. Canizares, "Probabilistic optimal power flow in electricity markets based on a two-point estimate method," *Power Systems, IEEE Transactions on*, vol. 21, no. 4, pp. 1883 –1893, 2006.
[11] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (p-opf) with gaussian and gamma distributions," *Power Systems, IEEE Transactions on*, vol. 20, no. 2, pp. 773 – 781, May 2005.
[12] R. Bo and F. Li, "Probabilistic lmp forecasting considering load uncertainty," *Power Systems, IEEE Transactions on*, vol. 24, no. 3, pp. 1279 –1289, Aug. 2009.
[13] A. Ott, "Experience with pjm market operation, system design, and implementation," *Power Systems, IEEE Transactions on*, vol. 18, no. 2, pp. 528 – 534, May 2003.
[14] G. Koop, *Bayesian Econometrics*. Wiley-Interscience, 2003.
[15] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *The Journal of Risk*, vol. 2, pp. 21–41, 2000.
[16] P. Huang and D. Subramanian, "Iterative estimation maximization for stochastic linear programs with conditional value-at-risk constraints," *Computational Management Science (forthcoming), http://www.optimization-online.org/DB_FILE/2008/04/1957.pdf*.
[17] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower's extensible optimal power flow architecture," in *Power Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, pp. 1 –7.

(a) Price

(b) Demand

Fig. 5. Price and demand dynamics: multiplicative model with risk-averse optimization.

The parameters of the posterior can be computed as

$$\bar{\nu} = \underline{\nu} + N$$
$$\bar{V} = \frac{1}{\underline{V}^{-1} + \sum_{n=1}^{N} x_n^2}$$
$$\bar{k} = \bar{V}(\underline{V}^{-1}\underline{k} + \hat{k}\sum_{n=1}^{N} x_n^2)$$
$$\bar{\nu}\bar{s}^2 = \underline{\nu}\underline{s}^2 + (N-1)s^2 + \frac{(\hat{k}-\underline{k})^2}{\underline{V} + 1/\sum_{n=1}^{N} x_n^2}$$

See [14] for details.