

# Probabilistic Forecast of Real-Time LMP via Multiparametric Programming

Yuting Ji, Robert J. Thomas, and Lang Tong

School of Electrical and Computer Engineering Cornell University, Ithaca, NY, USA.

Email: {yj246, rjt1,lt235}@cornell.edu

**Abstract**—The problem of short-term probabilistic forecast of real-time locational marginal price (LMP) is considered. A new forecast technique is proposed based on a multiparametric programming formulation that partitions the uncertainty parameter space into critical regions from which the conditional probability mass function of the real-time LMP is estimated using Monte Carlo techniques. The proposed methodology incorporates uncertainty models such as load and stochastic generation forecasts and system contingency models. With the use of offline computation of multiparametric linear programming, online computation cost is significantly reduced.

**Index Terms**—Locational marginal price (LMP), electricity price forecast, congestion forecast, probabilistic forecast, multiparametric programming.

## I. INTRODUCTION

As more renewable resources are integrated into the transmission system, and the power system operates closer to its capacity, congestion conditions become less predictable and LMPs more volatile.

The increased congestion and LMP uncertainties pose significant challenges to the operator and market participants, which motivates us to consider the problem of short-term forecast of real-time locational marginal price (LMP) in the presence of generation, demand, and operation uncertainties. A related problem is the forecast of transmission congestion—one of the main factors in the computation of LMP.

The benefit of LMP and congestion forecasts is twofold. For market participants, forecast of real-time prices is valuable in risk management, developing efficient bidding strategy, and demand side participation. The forecast price signal allows market participants to make adjustments in advance to ensure economic transactions.

For system operators, on the other hand, forecast of transmission congestion is important in congestion management, system planning, and operation. European transmission system operators, for instance, use Intraday Congestion Forecast (IDCF) to improve real-time security assessment [1] [2]. LMP forecast also alleviates congestion and facilitate demand response.

Currently, some system operators are providing real-time price forecasts. The Electric Reliability Council of Texas (ERCOT) [3] offers a 1 hour ahead real-time LMP forecast,

updated every 5 minutes. The Alberta Electric System Operator (AESO) [4] provides two short-term price forecasts with prediction horizons of 2 hours and 6 hours, respectively.

Most LMP forecast schemes fall into the category of point forecast. A point forecast algorithm gives a single quantity as the forecast value. For systems with highly random components and high level of uncertainties, point forecast is rarely accurate, and impacts of prediction error on decisions are difficult to quantify. A more attractive alternative is the probabilistic forecast that provides full characterization of the LMP distribution.

Significant technical challenges exist for probabilistic forecasting of real-time congestion and LMP. First, reasonably accurate models for real-time dispatch and LMP are needed. Second, real-time network operating conditions and uncertainties need to be incorporated. Finally, the forecast algorithm needs to be simple and scalable to sufficiently large systems. To this end, it is desirable to perform as much computation offline as possible. These challenges are daunting if the forecaster is merely a market participant without access to network operating conditions and network parameters. On the other hand, if it is the system operator performing the forecast, as in the case of ERCOT or AESO, the barrier to efficient and accurate forecast is significantly lowered.

### A. Summary of Contributions

In this paper, we consider the real-time LMP and congestion forecast problem from an operator perspective. We focus on *probabilistic forecast* that, at time  $t$ , the forecast algorithm provides the conditional distribution of the LMP vector and associated congestion status at time  $t + T$ . The main idea behind the proposed approach is the use of multiparametric program that partitions the uncertainty space into critical regions, and each region is attached to a unique LMP and a congestion pattern. Thus the problem of probabilistic forecast reduces to one of evaluating probabilities that the random parameter falls in a specific critical region. When loads or stochastic generations (as negative load) are random, load and generation forecast models are incorporated to generate probabilistic LMP and congestion forecasts. The proposed scheme can also incorporate custom specified reliability models of contingencies.

One of the key features of the proposed forecast methodology is to shift majority of computation offline, which signifi-

This work is supported in part by the National Science Foundation under Grant CNS-1135844 and a Grant from DoE Consortium for Electric Reliability Technology Solutions (CERTS).

cantly reduces the computation complexity when it is used to provide online rolling horizon prediction.

### B. Related Work

There is a substantial body of literature on LMP forecast; see [5] [6] and the reference therein. The majority of LMP forecast techniques deal with day-ahead LMP forecast, and the overwhelming majority focus on point forecast techniques [7] [8]. We highlight here the connection of the proposed approach to existing *probabilistic forecast methods*.

The idea of using multiparametric programming for real-time LMP forecast based on the partition of load space is new to our best knowledge. However, the idea of forecasting LMP probability distributions based on critical regions is explored in [9]. The authors of [9] considered the problem from a market participant perspective. Thus network parameters and operating conditions are not available to the forecaster, and critical regions cannot be constructed using a multiparametric program. It is also not easy to incorporate load/generation forecast models elsewhere in the network into local forecasts of LMP and congestion.

In [10], a probabilistic LMP forecast is proposed based on attaching a Gaussian distribution to a point estimate. The advantage of this approach is that it can be easily adopted by various point forecast methods. The disadvantage, on the other hand, is that the LMP is fundamentally discrete and the use of Gaussian distribution does not generate consistent forecasts. Similar to [9], this technique is limited to day-ahead forecast that does not utilize real-time operating conditions.

Another arena related to our work is short term transmission congestion forecast. There are a few papers focusing on probabilistic forecast [9] [11] [12]. The forecast problem is considered from the market participant side in [9] where inside system information are unavailable. On the other hand, the authors in [11] [12] proposed a similar approach from the system operator side. The forecasting technique proposed in [11] [12] relies on online Monte Carlo where the forecast algorithm solves an optimal power flow (OPF) problem for each Monte Carlo sample path, which carries a substantial computation cost. Such a high complexity algorithm is not scalable to large systems.

In terms of forecasting methodology, this paper is related to [13] with several key differences in the LMP model and forecasting techniques. Specifically, in this work, we consider an *ex-ante* LMP formulation whereas the formulation in [13] is based on an *ex-post* formulation. The techniques used here is also different. In particular, the approach in [13] is based on a non-homogeneous Markov chain model on a partition on the system state space. In this paper, in contrast, we focus on the partition directly on the load space. Instead of estimating transition probabilities from data in [13], we take advantage of probabilistic load forecasts.

This paper is organized as follows. Section II introduces the formulation of the ex-ante economic dispatch and the real-time LMP models. Section III provides the key theory basis of the proposed probabilistic forecast approach and Section

IV gives details of this algorithm. In Section V, the extension with contingency consideration is discussed. Numerical results are presented in Section VI and it follows the conclusion in Section VII.

## II. REAL-TIME EX-ANTE LMP MODEL

### A. Real-Time Economic Dispatch

In this paper, we consider an ex-ante real-time LMP model that arises from the real-time ex-ante economic dispatch. Specifically, the system operator solves a DC-OPF problem to find an optimal economic generation adjustment that meets the load forecast for the next interval and satisfies generation, transmission, and ramp constraints.

We describe here a standard ex-ante economic dispatch formulation of the real-time LMP model:

$$\begin{aligned} & \min_g c^T g \\ \text{subject to:} & \\ (\lambda) : & \mathbf{1}^T (g - d_{t+1|t}) = 0, \\ (\mu_+, \mu_-) : & -F_+ \leq \hat{A}(g - d_{t+1|t}) \leq F_+, \\ (\gamma_+, \gamma_-) : & g_- \leq g \leq g_+, \\ (\eta_+, \eta_-) : & \hat{g}_t - \Delta_- \leq g \leq \hat{g}_t + \Delta_+, \end{aligned} \quad (1)$$

where

$c$	vector of real-time offers;
$g$	vector of ex-ante dispatch at time $t + 1$ ;
$\hat{g}_t$	vector of generation estimate at time $t$ ;
$d_{t+1 t}$	vector of 1-step load forecast at time $t$ ;
$g_+/g_-$	max/min generator capacities;
$\Delta_-/\Delta_+$	upward/downward ramp limits;
$\hat{A}$	estimated shift factor matrix ;
$\lambda$	shadow price for the energy balance constraint;
$\mu_+/\mu_-$	shadow prices for transmission constraints;
$\gamma_+/\gamma_-$	shadow prices for capacity constraints;
$\eta_+/\eta_-$	shadow prices for ramp constraints.

In this model, we assume that each bus has a generator and a load, for simplicity. Note that the operating point  $\hat{g}_t$  and ramp limits are part of the linear program. By the ex-ante real-time LMP it means that the estimated system operating point  $\hat{g}_t$  and load forecast  $d_{t+1|t}$  are used in the computation of the economic dispatch and associated real-time prices.

### B. Real-time LMP Model

Assume that the shift factor matrix  $\hat{A}$  is constant over time. The Lagrangian of (1) is given by:

$$\begin{aligned} \mathcal{L} = & c^T g + \lambda \sum_i (g^{(i)} - d_{t+1|t}^{(i)}) \\ & + \sum_k \mu_+^{(k)} \sum_i \hat{A}_{ik} (g^{(i)} - d_{t+1|t}^{(i)}) \\ & - \sum_k \mu_-^{(k)} \sum_i \hat{A}_{ik} (g^{(i)} - d_{t+1|t}^{(i)}) \\ & + \sum_i \gamma_+^{(i)} (g^{(i)} - g_+^{(i)}) - \sum_i \gamma_-^{(i)} (g^{(i)} - g_+^{(i)}) \\ & + \sum_i \eta_+^{(i)} (g^{(i)} - \hat{g}_t^{(i)} - \Delta_+^{(i)}) \\ & - \sum_i \eta_-^{(i)} (g^{(i)} - \hat{g}_t^{(i)} + \Delta_-^{(i)}), \end{aligned}$$

where  $i$  is the bus index and  $k$  the transmission constraint index.

The partial derivative of the Lagrangian  $\mathcal{L}$  respect to the load vector  $d$  is given by:

$$\nabla_d \mathcal{L} = -\lambda \mathbf{1} - \hat{A}\mu_+ + \hat{A}\mu_-,$$

where  $\mathbf{1}$  is the vector of ones. By the Envelope Theorem, the real-time LMP  $\pi_t$  at time  $t$  can be expressed as the sum of the energy price and congestion prices

$$\pi_t = -\nabla_d \mathcal{L} = \lambda \mathbf{1} + \hat{A}\mu_+ - \hat{A}\mu_-. \quad (2)$$

### C. Certainty Equivalence Forecast of Real-time LMP

Given the economic dispatch and LMP models, we can introduce the certainty equivalence forecast of real-time LMP. The certainty equivalence forecaster simply treats the day-ahead schedule  $d_t^{\text{DA}}$  as being the actual realization  $d_t$  in real-time. The certainty equivalence forecast of LMP  $\pi_{t+T|t}$  at time  $t$  is equal to  $\pi_{t+T}$  using  $d_{t+T}^{\text{DA}}$  as the 1-step load forecast in (1). Note that the certainty equivalence forecast is a point prediction.

### D. Probabilistic Forecast of Real-time LMP

We now formulate the problem of *probabilistic* LMP forecast that, in contrast to the classical point forecast problem, aims to provide the probability distribution of LMP at a future time. In particular, given the estimated system operating point at time  $t$  and load and stochastic generation forecasts, the forecast probability mass function  $\hat{\pi}_{t+T|t}$  of LMP is computed.

The key to probabilistic LMP forecast is to capture spatial and temporal correlations and inherent system randomness. Spatial correlations among LMPs arise naturally from the optimization that governs the real-time economic dispatch. Temporal correlations, on the other hand, are results of that in load/generation forecasts and ramp constraints. The system randomness includes random occurrence of contingency and measurement noise.

## III. MULTIPARAMETRIC LINEAR PROGRAMMING

In this section, we provide the key theoretical foundation of the proposed probabilistic LMP forecast algorithm.

### A. Multiparametric Linear Programming Analysis

We adopt the formulation of multiparametric programming based on the right hand side multiparametric linear program (MLP-RHS) [14] [15]. To facilitate mathematical analysis, we rewrite constraints of linear program (1) in the following compact form:

$$Ug \leq b + Vd \quad (3)$$

where  $g$  is the optimization vector (generation dispatch),  $d$  is the vector of load/stochastic generation,  $U$ ,  $V$  and  $b$  are corresponding constant coefficients.

With the assumption of time invariant network topology, the uncertainty of real-time LMP only comes from random load/generation vector<sup>1</sup>  $d$ . Therefore, we treat load  $d$  as parameter vector.

<sup>1</sup>All other uncertainties are ignored at this point. Consideration of other variables is discussed in Section V.

Let  $\mathcal{D}$  be the feasible load region such that the ex-ante economic dispatch given by (1) has a finite optimal solution. Here, we want to characterize the relationship between the feasible load space  $\mathcal{D}$  and the real-time LMP  $\pi_t$ .

We recall following definitions in [14] for multiparametric linear programming analysis. Let  $J \triangleq \{1, \dots, m\}$  be the set of constraint indices in (3). For any  $I \subseteq J$ , let  $U_I$  and  $V_I$  be submatrices of  $U$  and  $V$ , respectively, consisting of rows indexed by  $I$ .

**Definition 1.** An *optimal partition* [14] of  $J$  associated with parameter  $d$  is the partition  $(I(d), I^C(d))$ , where

$$I(d) \triangleq \{i \in J | U_i g^*(d) = b + V_i d\},$$

$$I^C(d) \triangleq \{i \in J | U_i g^*(d) < b + V_i d\},$$

where  $g^*(d)$  is the optimal solution of (1) given  $d$ .

Given any load  $d$ , the optimal solution  $g^*(d)$  divides constraints into the binding set indexed by  $I(d)$  and the unbinding set indexed by  $I^C(d)$ . Such an optimal partition determines the congestion pattern (binding transmission constraints) and shadow prices can then be calculated. It follows that if the optimal partition is given, then the price is determined.

Furthermore, there exists a set of parameters that gives the same optimal partition, such a set is called *critical region*. More precisely, the definition of the critical region is given below.

**Definition 2.** For a given load vector  $d^* \in \mathcal{D}$ , let  $(I, I^C) \triangleq (I(d^*), I^C(d^*))$ , the *critical region* [14] related to the set of binding constraints  $I$  is defined as:

$$R_I \triangleq \{d \in \mathcal{D} | I(d) = I\}, \quad (4)$$

the set of all parameters  $d$  such that constraints indexed by  $I$  are binding at the optimum of linear program (1).

The correspondence of each critical region  $\mathcal{D}_i$  and each LMP vector  $\pi_i$  is summarized in the following lemma.

**Lemma 1.** A feasible load space  $\mathcal{D}$  can be partitioned into  $K$  critical regions in a unique way, *i.e.*

$$\mathcal{D} = \mathcal{D}_1 \uplus \mathcal{D}_2 \cdots \uplus \mathcal{D}_K, \quad (5)$$

where each critical region has the same LMP and the same congestion pattern, and critical regions  $\mathcal{D}_i$  are polyhedrons with boundaries assigned in consistent with shadow prices of the interior.

Proof: see [14].

Note that, given operation parameters,  $(U, b, V)$ , critical regions  $\mathcal{D}_i$  can be computed offline using techniques developed in [14].

## IV. REAL-TIME LMP PROBABILISTIC FORECAST

The basic idea of the proposed probabilistic forecast algorithm is illustrated in Figure 1 where the feasible load space is partitioned into critical regions and a realization of the load vector process forms a path on the space.

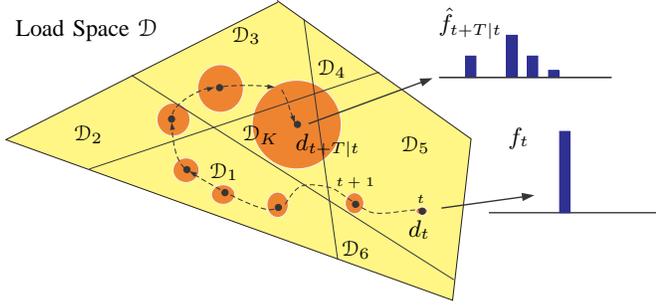


Figure 1: Geometric intuition of load random walk model.

Given the load realization  $d_t$  at time  $t$ , the optimal probabilistic forecast of real-time LMP at time  $t + T$  is given by conditional probabilities that load  $d_{t+T}$  falls in critical region  $\mathcal{D}_i$ , *i.e.*

$$\hat{f}_{t+T|t}(j) = \Pr[d_{t+T} \in \mathcal{D}_j | d_t]. \quad (6)$$

The computation of above conditional probabilities depends on probabilistic models of load. Such models can be obtained from either models of the load process or specific prediction methods used to generate load forecasts. Typically, such computation involves the use of Monte Carlo techniques. Approximations are necessary for high dimension problems.

As an illustration, we present a directional Gaussian random walk model for the load here. It should be noted that any statistical load model or prediction method can be applied. The purpose of using this model is to gain insights into the behavior of forecasting performance by taking advantage of some of analytically tractable properties.

Assume that load  $d_t$  evolves as a random walk process with incremental mean trajectory  $\Delta\mu_t = \mu_{t+1} - \mu_t = \mathbb{E}[d_{t+1}] - \mathbb{E}[d_t]$ , *i.e.*,

$$d_{t+1} = d_t + \Delta\mu_t + \Delta W_t, \quad (7)$$

where  $\Delta W_t \sim \mathcal{N}(0, \Sigma_t)$  is an independent Gaussian sequence. We then have

$$d_{t+T} = d_t + \sum_{i=0}^{T-1} \Delta\mu_{t+i} + W_{t+T}, \quad (8)$$

where  $W_{t+T} = \sum_{i=0}^{T-1} \Delta W_{t+i} \sim \mathcal{N}(0, \sum_{i=0}^{T-1} \Sigma_{t+i})$ .

Assume that the mean trajectory is the day ahead load forecast  $d_t^{\text{DA}}$ , then the random walk model becomes:

$$d_{t+1} = d_t + d_{t+1}^{\text{DA}} - d_t^{\text{DA}} + \Delta W_t. \quad (9)$$

In addition, we assume that at each time  $t$ , the load forecast noise of each bus is independent and its variance is constant, *i.e.*  $\Sigma_t = \Sigma$  and  $\Sigma$  is a diagonal matrix. Therefore, the actual load of time  $t + T$  can be written as:

$$d_{t+T} = d_t + d_{t+T}^{\text{DA}} - d_t^{\text{DA}} + W_{t+T}, \quad (10)$$

where  $W_{t+T} \sim \mathcal{N}(0, T\Sigma)$ .

The resulting conditional probability of  $d_{t+T}$  falling in critical region  $\mathcal{D}_i$  is:

$$\hat{f}_{t+T|t}(i) = \int_{\mathcal{D}_i} \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\{(x - \mu)^T \Sigma^{-1} (x - \mu)\} dx, \quad (11)$$

where  $\mu = d_t + d_{t+T}^{\text{DA}} - d_t^{\text{DA}}$ .

Specifically, in the one-dimensional case, the predictive probability of  $\pi_{t+T|t}$  equal to  $\pi_i$  is:

$$\hat{f}_{t+T|t}(i) = \Phi(\mathcal{D}_i^+) - \Phi(\mathcal{D}_i^-), \quad (12)$$

where  $\Phi(\cdot)$  is the cumulative distribution function of truncated Gaussian distribution  $\mathcal{N}(0, \Sigma)$  in the feasible load space  $\mathcal{D}$ ,  $\mathcal{D}_i^+$  and  $\mathcal{D}_i^-$  are upper and lower bounds of critical region  $\mathcal{D}_i$ , respectively. Note that the vector  $\hat{f}_{t+T|t}$  should be normalized as a distribution.

According to Lemma 1, the feasible load space  $\mathcal{D}$  can be uniquely partitioned into  $K$  critical regions, as illustrated conceptually in Figure 1. Each critical region  $\mathcal{D}_i$  corresponds to an LMP vector  $\pi_i$ . With an appropriate probabilistic load model, such as the random walk model described above, we can compute conditional probabilities in (6) via multiparametric programming analysis and Monte Carlo techniques.

The proposed algorithm can be summarized in following steps:

- 1) Use multiparametric linear programming analysis to determine the feasible load space  $\mathcal{D}$  and its critical regions.
- 2) Obtain LMP values and all correspondences between each pair of a LMP vector and a critical region.
- 3) Estimate the constant diagonal covariance matrix  $\Sigma$  of load forecast from historical data.
- 4) Compute conditional probabilities given in (11) with prediction horizon  $T$ .

First of all, we need to determine the feasible load space  $\mathcal{D}$  and its unique partition described in Lemma 1 respect to (1). This step can be achieved by the multiparametric programming analysis that computes all critical regions and associated dual variables. Note that the computation of critical regions can be done offline if ramping constraints are not active.

Next, for each critical region  $\mathcal{D}_i$ , we pick an arbitrary load vector  $d \in \mathcal{D}_i$ , and solve the linear program (1) to determine LMP values. Since each critical region corresponds to a unique LMP vector, we only need to solve this optimization problem once for each critical region.

Then, we use historical data of actual load and its day ahead forecast to estimate the constant covariance matrix  $\Sigma$ .

Finally, with all information obtained from previous steps and the newly available information  $d_t$  at current time  $t$ , we can predict the future LMP  $\pi_{t+T}$  via computing the conditional probability given by (11).

## V. DISCUSSION: CONTINGENCY CONSIDERATION

In this section, we extend the proposed forecast technique to incorporate contingency uncertainty models.

In the ex-ante economic dispatch model, contingencies can be viewed as a new set of coefficient values in (1). If there are transmission outages, values of the shift factor matrix  $\hat{A}$  and line limit  $F_+$  will change accordingly. Similarly, when unit outages happen, associated shift factors  $\hat{A}$ , generation capacities  $g_-, g_+$  and ramp limits  $\Delta_-, \Delta_+$  can be different from normal conditions.

For simplicity, we consider the N-1 contingency with the failure of a single element, which can be either a transmission line or a generator. Assume that a contingency  $i$  can happen with some probability  $p_i$  at any time in the prediction horizon. We also assume that once the contingency happens, it cannot be recovered within the prediction horizon, or a given period, for example, 24 hours. Then the probability that the contingency happens within the prediction horizon  $[t, t + T]$  is  $1 - (1 - p_i)^T$ .

To describe the extended algorithm, we restrict the single contingency to be a particular one with probability  $p$ . Multiple contingencies can be incorporated with different probability weights. Let  $\mathcal{M} = \{\hat{A}, F_+, g_+, g_-, \Delta_+, \Delta_-\}$  denote the set of constraint coefficients without contingency, and  $\tilde{\mathcal{M}}$  the set with a particular contingency.

We solve multiparametric linear programs with respect to  $\mathcal{M}$  and  $\tilde{\mathcal{M}}$  separately to obtain associated feasible load spaces  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$ . At each time  $t$ , if the contingency has not happened yet, the probabilistic LMP forecast is computed by

$$\hat{f}_{t+T|t} = (1 - p)^T \hat{f}_{t+T|t, \mathcal{D}} + [1 - (1 - p)^T] \hat{f}_{t+T|t, \tilde{\mathcal{D}}},$$

where  $\hat{f}_{t+T|t, \mathcal{D}}$  and  $\hat{f}_{t+T|t, \tilde{\mathcal{D}}}$  are conditional distributions of critical regions  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$ , respectively. Once the contingency is detected,  $\hat{f}_{t+T|t} = \hat{f}_{t+T|t, \tilde{\mathcal{D}}}$  by assumptions. To sum up, the probabilistic LMP forecast with contingency consideration is given by:

$$\hat{f}_{t+T|t} = \begin{cases} \hat{f}_{t+T|t, \tilde{\mathcal{D}}}, & \text{if contingency happened at time } t, \\ (1 - p)^T \hat{f}_{t+T|t, \mathcal{D}} \\ \quad + [1 - (1 - p)^T] \hat{f}_{t+T|t, \tilde{\mathcal{D}}}, & \text{otherwise.} \end{cases} \quad (13)$$

## VI. EVALUATION

In this section, we present some simulation results to compare performances of the proposed probabilistic forecast algorithm with the certainty equivalent forecaster. We first test our algorithm on a 3 bus system to gain insights into the behavior of the proposed algorithm under various scenarios, and then on the IEEE 14 bus system to prove its scalability.

Before presenting numerical results, we introduce a performance evaluation metric of probabilistic forecasts.

### A. Probabilistic Forecast Assessment

LMP is intrinsically a discrete random vector. The probabilistic forecast of such a random quantity belongs to the so-called categorical forecast, and its performance is measured by the consistency as well as the statistical concentration of the forecast. A standard metric is the Brier Score [16] that measures the average distance (2-norm) between the forecast

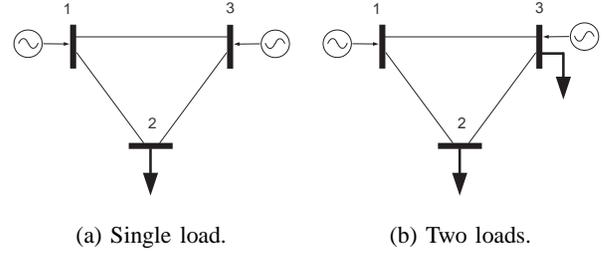


Figure 2: A 3 bus system.

distribution  $\hat{f}_{t+T|t}$  and the point mass distribution at the realized random variable  $\pi_{t+T}$ . Specifically

$$\text{BS}(\hat{f}_{t+T|t}) = \mathbb{E} \|\hat{f}_{t+T|t} - \delta(\pi_{t+T})\|^2, \quad (14)$$

where  $\hat{f}_{t+T|t}$  is the probability vector whose  $i$ th entry is given by  $\hat{f}_{t+T|t}(i) = \Pr(\pi = \pi_i)$ , and  $\delta(x)$  is the unit vector that is one at entry  $x$  and zero elsewhere. This score ranges from 0 to 2, where the larger the score, the worse the probabilistic prediction.

### B. Case Study: A 3 Bus System

The network topology of the 3 bus system is given in Figure 2. All three transmission lines are identical with thermal limits 100 MW. Generator at bus 1 has maximum capacity 140 MW with cost \$10. Generator at bus 3 has maximum capacity 200 MW with cost \$15. There are no minimum generation requirements.

1) *Scenario 1. A single load case:* We first consider a one dimensional load scenario, as shown in Figure 2a. Because the computation of conditional probabilities can be obtained analytically, this serves as a way of validating the behavior of the proposed algorithm.

The set up of this example is as follows. We generate a zigzag curve as the mean load trajectory varying from 120 MW to 180 MW which is used as the day ahead load forecast  $d_t^{\text{DA}}$  with 5-min interval, as shown in the upper part of Figure 3 with the right  $y$ -axis. To obtain actual load profiles, we simulate the random walk model described by (9) with an independent Gaussian sequence  $\Delta W_t \sim \mathcal{N}(0, \sigma^2)$ , by default,  $\sigma = 0.5\% \mathbb{E}[d_t^{\text{DA}}]$ .

We use a Matlab based multiparametric programming toolbox [17] to determine the feasible load space and its partition. The resulting critical regions<sup>2</sup> are:  $(0, 140]$ ,  $(140, 160]$  and  $(160, 200]$ . Note that boundary points, 140 MW and 160 MW, are assigned to critical regions with same LMP values. Load impacts on the prediction performance is first evaluated. From Figure 3, we can see that the proposed probabilistic forecast algorithm consistently outperforms the certainty equivalence predictor and longer prediction horizons result in worse forecasts. Furthermore, two interesting phenomena are observed: the presence of peaks and the increasing variance over time.

<sup>2</sup>We exclude the point  $d = 0$  for LMP value consistency in each critical region. For  $d = 0$ , values of LMP are zeros. But for critical region  $(0, 60]$ , LMPs are positive numbers.

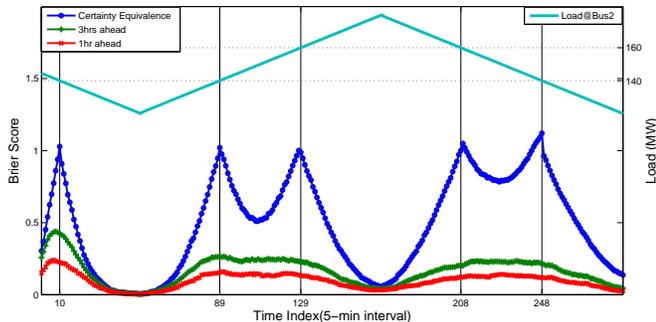


Figure 3: Impacts of load and prediction horizon in Scenario 1.

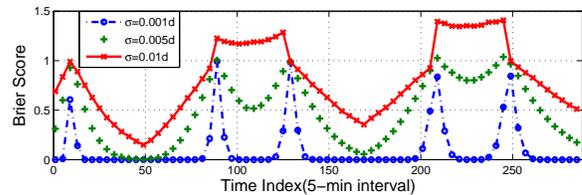
For the first phenomenon, we can easily check that all peaks coincidence with boundary points (indicated by  $x$ -axis ticks in Figure 3). When the mean trajectory load lies on the boundary, the probability of the actual load falling in either neighboring critical region is 0.5 where the value of the Brier Score using the certainty equivalence method becomes 1. In contrast, the probabilistic prediction is a conditional probability that incorporates all observed data. Hence, it optimally captures the direction of future loads.

The second phenomenon is more obvious on probabilistic forecasts, for instance, the 3 hours ahead probabilistic forecast curve. At time  $t = 10$ , the Gaussian distribution is much more concentrated on the mean than that at time  $t = 248$ . This is because that load follows a random walk model, so that the deviation from the mean trajectory grows, resulting in the bigger variance at a later time.

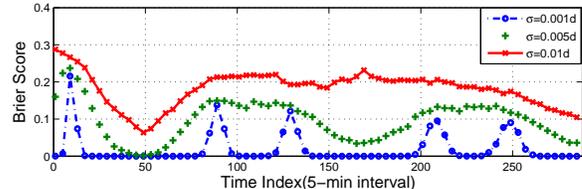
We then investigate the impact of the day-ahead load forecast quality. Three levels of load forecast error are considered:  $\sigma = 0.1\% \mathbb{E}[d_t^{DA}]$ ,  $\sigma = 0.5\% \mathbb{E}[d_t^{DA}]$  and  $\sigma = 1\% \mathbb{E}[d_t^{DA}]$ . The resulting performances of the certainty equivalence predictor, the probabilistic forecasts with prediction horizon of 1 hour and 3 hours are shown in subfigures of Figure 4, respectively. From the comparison between subfigures, we can see that all algorithms perform worse with bigger load forecast errors. In each subfigure, the certainty equivalent prediction appears to be more sensitive to the load error than probabilistic forecasts.

2) *Scenario 2. A two dimensional load case:* We then consider a two dimensional load vector scenario where the system network setting is shown in Figure 2b. To generate a reasonable mean trajectory of load profile, we draw an arc crossing all three critical regions indicated as the arc in Figure 5. For this particular load forecast profile, each bus follows a sinusoidal wave as shown in Figure 6. Note that boundary points of neighboring critical regions are highlighted on  $x$ -axis ticks.

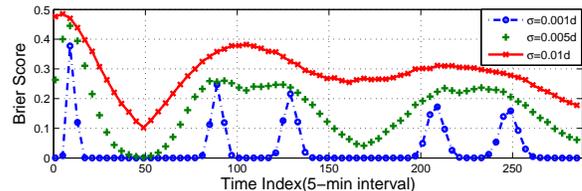
Actual load profiles are generated from the random walk model in (9) with two independent Gaussian sequences. The standard deviation of each bus load is assumed to be .5% of its mean. A Monte Carlo method is used to estimate conditional probabilities of critical regions in (11).



(a) Certainty equivalent prediction.



(b) 1 hour ahead probabilistic forecast.



(c) 3 hours ahead probabilistic forecast.

Figure 4: Impact of load forecast quality in Scenario 1.

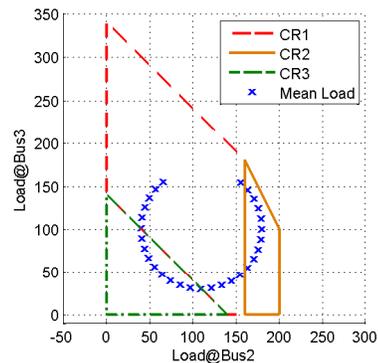


Figure 5: The feasible load space, critical regions and a mean load trajectory in Scenario 2.

We conduct the same experiment as Scenario 1 in this case. As shown in Figure 6, in addition to lower Brier Scores, probabilistic forecasts are also less sensitive at boundary points. Similar phenomenon of increasing distribution variance is observed.

3) *Scenario 3. A single contingency case:* Finally, the extended algorithm with contingency consideration is evaluated. In this setting, we consider a unit outage with a partial loss of generation capacity. Assume that the maximum capacity of the generator located at bus 1 can be reduced to 70 MW from 140 MW with probability  $p$  at any time. Other settings are the same as the first scenario. The feasible load space with this particular contingency becomes  $(0, 185]$  and associated critical regions become  $(0, 70]$  and  $(70, 185]$ . Therefore,

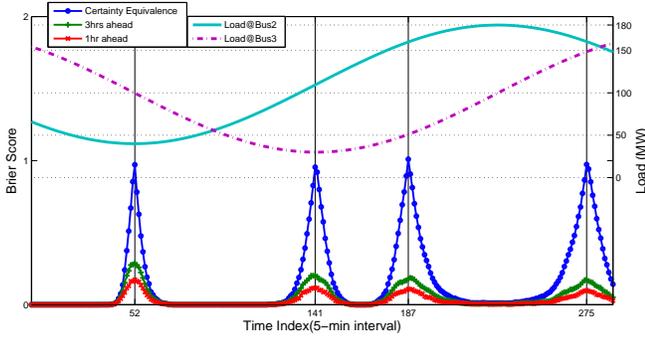


Figure 6: Impacts of load and prediction horizon in Scenario 2.

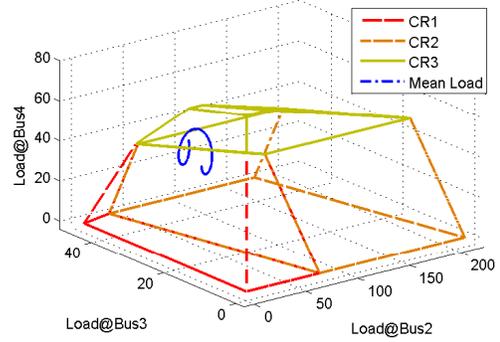


Figure 8: The feasible load space, critical regions and a mean load trajectory in Scenario 3.

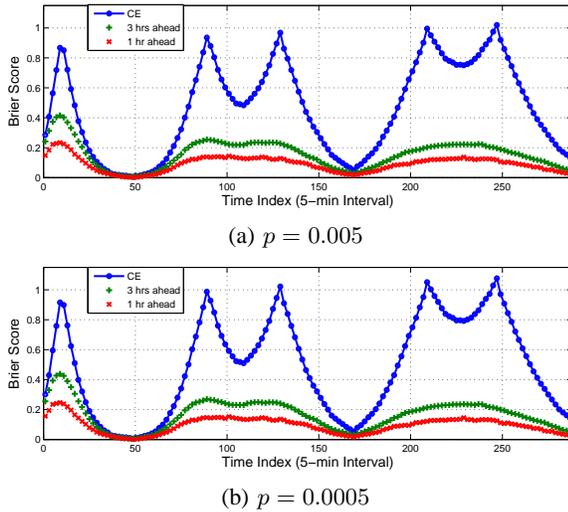


Figure 7: Impact of contingency occurrence probability  $p$  in Scenario 3.

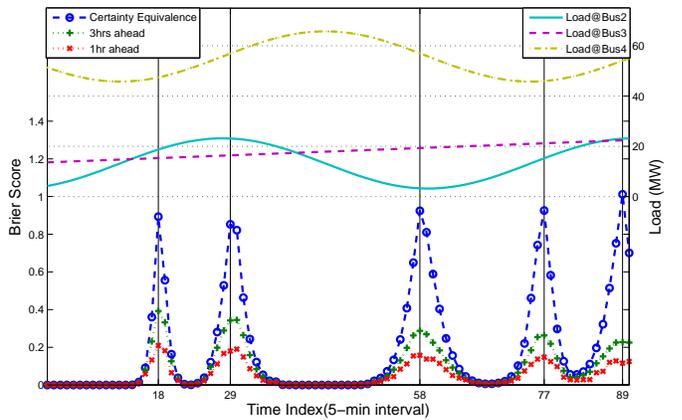


Figure 9: Impacts of load and prediction horizon in Scenario 3.

there are five critical regions that a future load can fall in:  $(0, 140]$ ,  $(140, 160]$ ,  $(160, 200]$ ,  $(0, 70]^*$ , and  $(70, 185]^*$ , where critical regions with the presence of contingency are labeled by “\*”. We compare the extended algorithm described in Section V with the extended certainty equivalence prediction, which also takes the contingency probability into consideration. Since the outage frequency is very low in practice, we choose  $p = 0.005$  and  $p = 0.0005$ . Results are shown in Figure 7. Performances are similar to that of the normal scenario, as shown in Figure 3. However, the relatively high outage ratio performs better than the lower ratio. This counterintuitive phenomenon is caused by the particular load range  $([120, 180])$  we chose. Once the contingency happens, all real-time actual loads will fall in the fifth critical region, *i.e.*,  $(70, 180]$ , hence, the uncertainty is considerably reduced.

### C. Case Study: IEEE 14 Bus System

We use the IEEE 14 Bus system with line limit parameters given in [18]. In this case, we only consider generators located at bus 1 and 2 as online units since all other generators are

not designed for providing active power. The load parameter vector consists of loads at bus 2, 3 and 4. The feasible space, its critical regions and the mean trajectory are shown in Figure. 8. Performances of the certainty equivalent prediction and probabilistic forecasts are shown in Figure 9.

## VII. CONCLUSION

This paper presents a new methodology for the short-term forecast of real-time LMP. The key idea is the exploitation of stochastic models to load uncertainty by the use of multiparametric linear programming analysis and incorporating online measurements.

The proposed technique does have several issues that need to be addressed in the future. We have not discussed how to compute conditional probabilities of future load in critical regions. In principle, Monte Carlo techniques can be applied, but the accuracy of such techniques is limited by the number of samples generated; a more sophisticated sampling technique is particularly useful in obtaining accurate estimates. Another issue arises from ramp constraints. Although this paper provides a heuristic algorithm to incorporate ramp constraints,

the feasibility can be an issue in practice due to load forecast error. A more accurate forecast algorithm of feasible space and critical regions should be explored.

As a general forecast technique, the proposed algorithm can be tailored to take advantage of various forecasts which can be very useful in practice, especially for the integration of renewable resources.

## REFERENCES

- [1] TSC, Intraday Congestion Forecast (IDCF), <http://www.amprion.net/pressemitteilung-126>.
- [2] PSE S.A., <http://www.psi.de/en/psi-pressevents/releases-archive/article/article/psi-receives-contract-for-implementing-intraday-congestion-forecast-for-polish-transmission-system-o/>.
- [3] ERCOT, Wholesale Price Forecast Tool, [http://www.ercot.com/news/press\\_releases/show/26244](http://www.ercot.com/news/press_releases/show/26244).
- [4] AESO, Pool Price Forecast Calculation Methodology, [http://www.aeso.ca/downloads/Price\\_Forecast\\_Calculation\\_March\\_8\\_2011.pdf](http://www.aeso.ca/downloads/Price_Forecast_Calculation_March_8_2011.pdf).
- [5] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *International Journal of Electrical Power and Energy Systems*, vol. 31, no. 1, pp. 13 – 22, 2009.
- [6] G. Klaeboe, A.L.Eriksrud, and S. Fleten, "Benchmarking time series based forecasting models for electricity balancing market prices," *International Journal of Energy System*, 2013.
- [7] F. Nogales, J.Contreras, A. Conejo, and R. Espfnola, "Forecasting next-day electricity prices by time series models," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342–348, 2002.
- [8] A. Conejo, J. Contreras, R. Espinola, and M. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," *International Journal of Forecast*, vol. 21, pp. 435–462, 2005.
- [9] Q. Zhou, L. Tsatsion, and C. Liu, "Short-term congestion forecasting in wholesale power markets," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2185–2196, 2011.
- [10] R. Bo and F. Li, "Probabilistic LMP Forecasting Considering Load Uncertainty," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1279–1289, 2009.
- [11] L. Min, S. T. Lee, P. Zhang, V. Rose, and J. Cole, "Short-term probabilistic transmission congestion forecasting," in *Proc. of the 3rd International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, 2008, pp. 764–770.
- [12] Electric Power Research Institute, Probabilistic Transmission Congestion Forecasting, [http://uc-ciee.org/downloads/PTCF\\_Final\\_Report.pdf](http://uc-ciee.org/downloads/PTCF_Final_Report.pdf).
- [13] Y. Ji, J. Kim, R. J. Thomas, and L. Tong, "Forecasting Real-time Locational Marginal Price: A State Space Approach," in *Proc. of the 47th Asilomar Conference on Signals, Systems, and Computers*, 2013, pp. 379–383.
- [14] F. Borrelli, A. Bemporad, and M. Morari, "Geometric algorithm for multiparametric linear programming," *Journal of Optimization Theory and Applications*, vol. 118, no. 3, pp. 515–540, 2003.
- [15] T. Gal and J. Nedoma, "Multiparametric linear programming," *Management Science*, vol. 18, pp. 406–442, 1972.
- [16] G. W. Brier, "Verification of forecasts expressed in terms of probability," *Monthly Weather Review*, vol. 78, pp. 1–3, 1950.
- [17] M. Herceg, M. Kvasnica, C. Jones, and M. Morari, "Multi-Parametric Toolbox 3.0," in *Proc. of the European Control Conference*, Zürich, Switzerland, July 17-19 2013, pp. 502–510, <http://control.ee.ethz.ch/~mpt>.
- [18] P. Venkatesh, R. Gnanadass, and N. P. Padhy, "Comparison and application of evolutionary programming techniques to combined economic emission dispatch with line flow constraints," *IEEE Transactions on Power Systems*, vol. 18, no. 2, 2003.