Using a System Model to Decompose the Effects of Influential Factors on Locational Marginal Prices

Lizhi Wang and Mainak Mazumdar

***Abstract*—Locational marginal prices (LMPs) are influenced by various factors such as load uncertainty, thermal limit, capacity reserve, and market power. We build a system model to quantitatively analyze the effects of these individual factors and their interactions on the mean and standard deviation of the LMPs, assuming that the factors are either active or inactive. According to the sensitivity analysis results from convex quadratic programming, LMPs are piecewise linear functions of demand variation, which is a key property used in the computation. This paper attempts to answer the following types of questions: (1) To what extent does market power raise the LMPs above the marginal cost? (2) What effects will generation or transmission capacity expansion have on relieving high LMPs under high demand scenarios? An IEEE 30-bus test system is used as an example to demonstrate our approach.**

***Index Terms*—Capacity expansion, load flow analysis, locational marginal price (LMP), power system planning, probability, quadratic programming, sensitivity analysis.**

I. Introduction

THE formation of locational marginal prices (LMPs) is a location-dependent stochastic process, which is driven by a combination of various factors [1]. Valenzuela and Mazumdar [2] categorize these factors into physical factors (which include production cost, load, generation availability, unit commitment, and transmission constraints) and economic factors (which include strategic bidding and load elasticity). One approach to analyzing these factors is to derive analytical expressions for the sensitivity of LMPs with respect to the parameters of the optimal power flow models which determine those prices [3], including the sensitivity with respect to bid parameters. A similar approach has been used to identify market power in the energy market [4]. Many of these factors are stochastic by nature, and they jointly affect the LMP probability distribution. It is therefore hard to determine from historical data the sole or interactive contributions of the individual factors (see [5]–[7] for examples of empirical studies of LMPs). However, such information is important and instructive in various ways. The exercise of market power, for example, has been a big concern since the beginning of electricity market deregulation; thus, it will be useful to quantitatively distinguish the sole contribution of market power from that of other factors in raising the LMPs above marginal cost. Another use of this information is to accurately evaluate the effects on LMPs when generation or transmission capacity expansion plans are being made. The objective of this paper is to build a system model to analyze and decompose the effects of various influential factors on the LMP probability distributions, assuming that the factors are binary variables (inactive or active). In this task, the property that the LMP is a piecewise linear function of demand variation turns out to be handy.

For ease of exposition, only four of the most important factors are considered: load uncertainty, thermal limit, capacity reserve, and market power, numbered by 1, 2, 3, and 4, respectively. (The effects of other factors, e.g., fuel price changes, are not being studied here.) Our system model will analyze the contribution of each single and combination of factors to the means and standard deviations of LMPs at different nodes. The mean of LMPs measures the long-term average level of prices, while the standard deviation is a measure of LMP variability in the same unit with LMPs ($/MWh). The statistical models for node n in hour t are in the following two postulated linear forms:

$$\begin{aligned}\mu^n_{\mathrm{LMP}}(t) = {} & \mu^n_0(t) + \mu^n_1(t) + \mu^n_2(t) + \mu^n_3(t) + \mu^n_4(t) \\ & + \mu^n_{12}(t) + \mu^n_{13}(t) + \mu^n_{14}(t) + \mu^n_{23}(t) + \mu^n_{24}(t) \\ & + \mu^n_{34}(t) + \mu^n_{123}(t) + \mu^n_{124}(t) + \mu^n_{134}(t) \\ & + \mu^n_{234}(t) + \mu^n_{1234}(t) \qquad (1)\\ \sigma^n_{\mathrm{LMP}}(t) = {} & \sigma^n_0(t) + \sigma^n_1(t) + \sigma^n_2(t) + \sigma^n_3(t) + \sigma^n_4(t) \\ & + \sigma^n_{12}(t) + \sigma^n_{13}(t) + \sigma^n_{14}(t) + \sigma^n_{23}(t) + \sigma^n_{24}(t) \\ & + \sigma^n_{34}(t) + \sigma^n_{123}(t) + \sigma^n_{124}(t) + \sigma^n_{134}(t) \\ & + \sigma^n_{234}(t) + \sigma^n_{1234}(t). \qquad (2)\end{aligned}$$

Here, $\mu^n_{\mathrm{LMP}}(t)$ and $\sigma^n_{\mathrm{LMP}}(t)$ are the realized mean and standard deviation of LMPs at node n in hour t, respectively; $\mu^n_0(t)$ and $\sigma^n_0(t)$ are, respectively, the mean and standard deviation given none of the four factors' presence; for $i, j = 1, 2, 3, 4$, $\mu^n_i(t)$ is the sole contribution of factor i to the mean of LMPs at node n in addition to $\mu^n_0(t)$, and $\sigma^n_{ij}(t)$ is the contribution resulting from (exclusively) the interaction between factors i and j to the standard deviation of LMPs at node n in addition to $\sigma^n_0(t)$. This is illustrated in Fig. 1 for a three-factor case. The rectangular area as a result of three interacting factors is decomposed to $\mu_0 + \mu_i + \mu_j + \mu_k + \mu_{ij} + \mu_{jk} + \mu_{ik} + \mu_{ijk}$. We will refer to the coefficients in (1) and (2) as *impact coefficients*, and our primary objective is to use the system model to obtain these impact coefficients. We point out here that the validity of this model is

Manuscript received October 9, 2006; revised June 12, 2007. This work was supported by the National Science Foundation under Grant ECS-0245355. Paper no. TPWRS-00691-2006.

L. Wang is with the Department of Industrial and Manufacturing Systems Engineering and the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: lzwang@iastate.edu).

M. Mazumdar is with the Department of Industrial Engineering, University of Pittsburgh, Pittsburgh, PA 15260 USA (e-mail: mmazumd@engr.pitt.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2007.907121

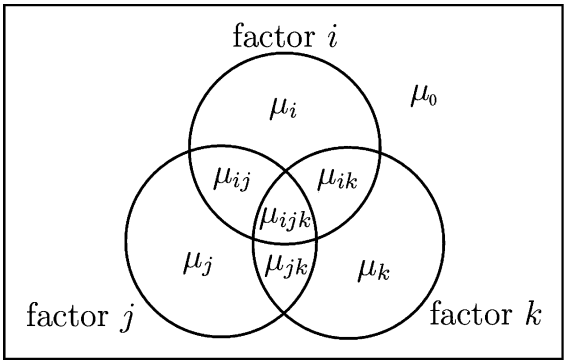

Fig. 1. Decomposition of contribution by factors i, j, and k.

based on the binary properties of the factors (either active or inactive); a different model would be needed if the factors were instead assumed to be continuous.

Section II describes the system model and the four factors in more detail. Section III first defines the inputs and outputs of the system model in III-A; and then, III-B describes the derivation of outputs from inputs by utilizing the piecewise linear property of LMP as a function of demand variation; the approach used in obtaining impact coefficients from input and output is given in III-C. A numerical example on the IEEE 30-bus network is given in Section IV, where IV-A gives the data of the test system, IV-B and IV-C present and interpret the time-averaged impact coefficients and their time variation, respectively. Use of the system model in evaluating transmission and generation capacity expansion plans is discussed and illustrated in Section V. Section VI concludes this paper.

II. System Model

A. Transmission Network

A set of nodes $\mathcal{N}$ is connected by a set of transmission lines $\mathcal{L}$. The sets of nodes with demand for and supply of power are denoted by $\mathcal{D}$ and $\mathcal{S}$, respectively. Depending on whether or not there is demand for or supply of power, any node in $\mathcal{N}$ could belong to either $\mathcal{D}$ or $\mathcal{S}$, or both, or neither.

B. Load Uncertainty

For a certain length of period $t = 1, 2, \ldots, T$, demand is assumed to be inelastic: $d_{n,t}(1 + \epsilon_t)$, $\forall n \in \mathcal{D}, t = 1, 2, \ldots, T$, where $d_{n,t}$ (in MW) is the nominal load at node n in hour t, while ϵ_t is a random variable, representing the demand uncertainty in percentage of $d_{n,t}$. Notice that for a given t, the load uncertainty ϵ_t is assumed to be the same at all nodes, which means that demands at all nodes are perfectly correlated, so that they increase or decrease universally by the same percentage.

In our system model, $d_{n,t}$'s are assumed to be known constants, and ϵ_t's are assumed to have known probability distributions, which can be obtained from the historical load data.

C. Thermal Limit

A dc lossless load flow model is used here, which has been found to be a good approximation to the more accurate ac load flow model when thermal limits are the primary concern [8], [9].

Denote by z_n, H, and T_l the net injection at node n, power transfer distribution factors (PTDF) matrix, and capacity of transmission line l, respectively. Net injection is the total power flow going into a node less the total power flow going out of it. PTDF matrix gives the linear relation between net injection at each node and power flow through each line. For all $l \in \mathcal{L}$, $|\sum_{n\in\mathcal{N}} H_{l,n} z_n|$ calculates the magnitude of the power flow through line l. If a transmission line's thermal limit is exceeded for a significant length of time, conductors sag or can be damaged by excessive heating, and the probability of short-circuiting with the ground increases. Therefore, the transmission constraints require that power flow going through any transmission line in either direction must be within the capacity

$$\sum_{n\in\mathcal{N}} H_{l,n} z_n \le T_l, \ \forall l \in \mathcal{L}$$
$$-\sum_{n\in\mathcal{N}} H_{l,n} z_n \le T_l, \forall l \in \mathcal{L}.$$

D. Capacity Reserve

As part of the ancillary services, certain amounts of generation and transmission capacities are kept in reserve to be able to reestablish the balance between load and generation in the event of a contingency. However, obtaining the exact amount of reserve capacity that is "optimal" for all stakeholders is a complex problem, and the solution may vary depending on the perspective chosen. The N-1 criterion, for example, requires that the reserve level should be sufficient to counter the loss of any single component (generator or transmission line). On the other hand, [10] and [11] simply derate the capacities by the forced outage rates to account for contingencies. As opposed to the above deterministic criteria, stochastic criteria have also been proposed [12]–[14], where transmission line reliability is taken into account in determining the reserve level.

In our model, as an illustration, we use what we refer to as the 90% criterion. This criterion is to require 10% reserve capacities of all generators and transmission lines.

In considering the capacity reserve factor, our purpose is not to study the LMP probability distribution under contingencies but simply to address the fact that when system security is taken into account, the system operator would be restricted from fully utilizing all the available generation and transmission capacities.

E. Market Power

Following [15], we assume that there is a single generator at each supply node (we will refer to the generator at node n as generator n, which should not give arise to any confusion), having a marginal cost function

$$q_n \mapsto a_n + b_n q_n, \quad \forall n \in \mathcal{S}$$

where q_n (in MWh) is the quantity of power generation at node n, a_n (in \$/MWh) and b_n (in $\$/(\mathrm{MWh})^2$) are constant param-

eters. The generating firms submit a linear supply function for each of their generators

$$q_n \mapsto \alpha_n + b_n q_n, \quad \forall n \in \mathcal{S}$$

and they exercise their market power by strategically submitting α_n's that may be different from a_n's to maximize their profits, and they are simply assumed not to manipulate on b_n's.

Instead of solving α_n's using game theoretic models as in [15], we assume them to be known parameters, and we attempt to find out quantitatively the effects of their numerical differences from the a_n's. We also assume that the supply functions stay the same for the entire time horizon.

F. Market Clearing

The electricity market is cleared each hour using the following economic dispatch:

$$\min_{q,z} \quad \sum_{n\in\mathcal{S}} \left(\alpha_n q_n + \frac{1}{2} b_n q_n^2\right) \tag{3}$$

$$\text{s.t.} \quad q_n - z_n = d_{n,t}(1+\epsilon_t), \quad (p_{n,t}) \forall n \in \mathcal{N} \tag{4}$$

$$\sum_{n\in\mathcal{N}} H_{l,n} z_n \le 0.9 T_l, \quad \forall l \in \mathcal{L} \tag{5}$$

$$-\sum_{n\in\mathcal{N}} H_{l,n} z_n \le 0.9 T_l, \quad \forall l \in \mathcal{L} \tag{6}$$

$$\sum_{n\in\mathcal{N}} z_n = 0 \tag{7}$$

$$0 \le q_n \le 0.9\overline{Q_n}, \quad \forall n \in \mathcal{S}. \tag{8}$$

Here, $\overline{Q_n}$ is the capacity of generator n, and dual variable $p_{n,t}$ is the LMP at node n in hour t. The objective function (3) is to minimize the generation cost (using the firm submitted supply functions); (4) comes from the definition of net injection, and the dual variable $p_{n,t}$ calculates the marginal cost of serving unit increment of demand at node n in hour t, which is consistent with the definition of LMP; (5) and (6) are transmission constraints under the 90% criterion; (7) is the balancing property of a network; and (8) is the generation capacity limit under the 90% criterion.

The economic dispatch (3)–(8) is a convex quadratic program and is generally easy to solve in a given hour for a given value of ϵ_t. It is assumed that generation and transmission capacities are sufficient to serve demand at all scenarios, so that an optimal solution to (3)–(8) always exists.

G. Other Factors

The system model can also be used to analyze the effects of other factors. For example, we can quantitatively examine the effects of introducing new generators and/or expanding the capacities of existing transmission lines. This example will be illustrated in Section V, where factors 1, 2, 3, and 4 are all assumed to be active, and four new factors 5, 6, 7, and 8 are introduced.

TABLE I
INTERPRETATION OF INPUT VARIABLES

i	$x_i = 1$	$x_i = 0$
1	Demand uncertainty is active: ϵ_t is a random variable, thus $p_{n,t}(\epsilon_t)$ will be a random variable as well.	Demand has no uncertainty: ϵ_t is a constant 0, and $p_{n,t}$ will also be a constant.
2	Thermal limit is active: congestion may occur when demands are high.	Thermal limit is ignored: constraints (5) and (6) are ignored and there will be no congestion.
3	The 90% criterion is used as in constraints (5), (6) and (8).	The 90% criterion is not used in constraints (5), (6) and (8).
4	Market power is exercised to a given extent (see Table II) as in the objective (3).	Market power is not exercised: firms honestly submit marginal cost functions as supply functions, thus a_n instead of α_n will be used in the objective (3).

III. Determination of Impact Coefficients

A. Input and Output of the System Model

We define four binary input variables x_1, x_2, x_3, and x_4 to represent the presence of load uncertainty, thermal limit, capacity reserve, and market power, respectively. Table I gives the interpretation of these input variables. In reality, all the factors are active. In the system model, however, some of the factors need to be assumed absent so that the difference it makes can be obtained and used to calculate the impact coefficients.

For a given set of input variables $\mathbf{x} = \{x_1, x_2, x_3, x_4\}$, the output variables of the system model at node n are $p^*_{n,t}(\mathbf{x};\epsilon_t)$, $\mu^n_t(\mathbf{x})$, and $\sigma^n_t(\mathbf{x})$. Here $p^*_{n,t}(\mathbf{x};\epsilon_t)$ is the LMP at node n in hour t, which results from the optimal (dual) solution to the economic dispatch (3)–(8) and is also a function of ϵ_t; $\mu^n_t(\mathbf{x})$ and $\sigma^n_t(\mathbf{x})$ are, respectively, the mean and standard deviation of $p^*_{n,t}(\mathbf{x};\epsilon_t)$. Let the probability distribution function (pdf) of ϵ_t be $f_t(\cdot)$, then

$$\begin{aligned}\mu^n_t(\mathbf{x}) &= E_{\epsilon_t}\left[p^*_{n,t}(\mathbf{x};\epsilon_t)\right] \\ &= \int f_t(\epsilon_t) p^*_{n,t}(\mathbf{x};\epsilon_t) d\epsilon_t, \text{ and} \\ \sigma^n_t(\mathbf{x}) &= \sqrt{V_{\epsilon_t}\left[p^*_{n,t}(\mathbf{x};\epsilon_t)\right]} \\ &= \sqrt{\int f_t(\epsilon_t)\left[p^*_{n,t}(\mathbf{x};\epsilon_t) - \mu^n_t(\mathbf{x})\right]^2 d\epsilon_t}\end{aligned}.$$

B. Deriving Output From Input

This subsection describes the algorithm for deriving $p^*_{n,t}(\mathbf{x};\epsilon_t)$ as a function of ϵ_t from the system model for a given $\mathbf{x}$.

When $x_1 = 0$, the economic dispatch (3)–(8) only needs to be solved once to obtain the output for a given input. When $x_1 = 1$, regardless of the other input variables, the economic dispatch (3)–(8) can be represented by the following standard form parametric convex quadratic program (QP_ϵ):

$$\min_x \left\{ c^\top x + \frac{1}{2} x^\top Q x : Ax = b + \epsilon\Delta b, x \ge 0 \right\}$$

where $c \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$, and $b \in \mathbb{R}^m$ are constants, Q is a positive semi-definite matrix, $\Delta b \in \mathbb{R}^m$ is a given direction of variation, and ϵ is a scalar parameter. In the economic dispatch context, $\Delta b = [d_{1,t}, \cdots, d_{|\mathcal{N}|,t}, 0, \cdots, 0]^\top$. It is well known that the optimal solution $x^*(\epsilon)$ is a piecewise linear function of ϵ. Reference [16] gives an algorithm that analytically calculates the break points and the functional form of each segment. We can use that algorithm to obtain the piecewise linear function $p^*_{n,t}(1, x_2, x_3, x_4; \epsilon_t)$ and then calculate other outputs.

We present below the algorithm that we use in our computational experiments to obtain the break points of ϵ_t. The functional form of $p^*_{n,t}(1, x_2, x_3, x_4; \epsilon_t)$ within each segment can be calculated by solving at the lower and upper limit break points and then connecting them with a straight line. Our algorithm adopts the basic ideas from [16] but is computationally more robust.

We first review some preliminaries of parametric convex quadratic programming. Let (QD_ϵ) be the Wolfe dual of (QP_ϵ)

$$\max_{x,y,s}\left\{(b+\epsilon\Delta b)^\top y - \frac{1}{2}x^\top Qx : A^\top y + s - Qx = c, s \geq 0\right\}.$$

For a given ϵ, the tripartition $\pi(\epsilon) = \{\mathcal{B}(\epsilon), \mathcal{N}(\epsilon), \mathcal{T}(\epsilon)\}$ is defined as

$$\begin{aligned}\mathcal{B}(\epsilon) &= \{i : x_i > 0 \text{ for an optimal solution } (x(\epsilon), y(\epsilon), s(\epsilon))\}\\ \mathcal{N}(\epsilon) &= \{i : s_i > 0 \text{ for an optimal solution } (x(\epsilon), y(\epsilon), s(\epsilon))\}\\ \mathcal{T}(\epsilon) &= \{1, \ldots, n\} \setminus \{\mathcal{B}(\epsilon) \cup \mathcal{N}(\epsilon)\}.\end{aligned}$$

Define a maximal complementary solution $(x^*(\epsilon), y^*(\epsilon), s^*(\epsilon))$ as an optimal solution such that

$$x^*_i(\epsilon) > 0 \Leftrightarrow i \in \mathcal{B}(\epsilon) \text{ and } s^*_i(\epsilon) > 0 \Leftrightarrow i \in \mathcal{N}(\epsilon).$$

It has been shown in [17] and [18] that a maximal complementary solution can be obtained by solving (QP_ϵ) or (QD_ϵ) using interior point methods.

We present the algorithm as pseudo codes of two functions: `main_function` and `sub_function`. The function inputs of `main_function` are coefficients of (QP_ϵ) and (QD_ϵ), and the function output is the set of break points of ϵ within the entire range that (QP_ϵ) and (QD_ϵ) remain feasible. The function inputs of `sub_function` are coefficients of (QP_ϵ) and (QD_ϵ) and a range $(\underline{\epsilon}, \overline{\epsilon})$, and the function output is the set of break points of ϵ within $(\underline{\epsilon}, \overline{\epsilon})$.

$\Upsilon = \texttt{main\_function}(A, b, \Delta b, c, Q)$

{**Step 1**: Calculate (L_ϵ, U_ϵ), which is the entire range of ϵ that (QP_ϵ) and (QD_ϵ) are feasible:

$$\begin{aligned}L_\epsilon = \min_{\epsilon,x,y,s}\{&\epsilon : Ax - \Delta b\epsilon = b, A^\top y + s - Qx = c\\ &x \geq 0, s \geq 0\}\\ U_\epsilon = \max_{\epsilon,x,y,s}\{&\epsilon : Ax - \Delta b\epsilon = b, A^\top y + s - Qx = c\\ &x \geq 0, s \geq 0\}.\end{aligned}$$

Step 2: Call $\Upsilon = \texttt{sub\_function}(A, b, \Delta b, c, Q, L_\epsilon, U_\epsilon)$, and return Υ.

}

$\Upsilon = \texttt{sub\_function}(A, b, \Delta b, c, Q, \underline{\epsilon}, \overline{\epsilon})$

{**Step 1**: If $(\overline{\epsilon} - \underline{\epsilon})$ is sufficiently small (or both $\overline{\epsilon}$ and $\underline{\epsilon}$ are infinity with same sign), then return $\Upsilon = (1/2)(\overline{\epsilon} + \underline{\epsilon})$. Otherwise, set the initial point $\epsilon^0 = (1/2)(\overline{\epsilon} + \underline{\epsilon})$ and continue.

Step 2: Obtain the tripartition

$$\pi(\epsilon^0) = \{\mathcal{B}(\epsilon^0), \mathcal{N}(\epsilon^0), \mathcal{T}(\epsilon^0)\}$$

by solving $(\mathrm{QP}_{\epsilon^0})$ and $(\mathrm{QD}_{\epsilon^0})$ using an interior point method and obtaining a maximal complementary solution $(x^*(\epsilon^0), y^*(\epsilon^0), s^*(\epsilon^0))$.

Step 3: Calculate $(l_{\epsilon^0}, u_{\epsilon^0})$, which is the range of ϵ that (QP_ϵ) and (QD_ϵ) have the same tripartition as $\pi(\epsilon^0)$:

$$\begin{aligned}l_{\epsilon^0} = \min_{\epsilon,x,y,s}\{&\epsilon : Ax - \Delta b\epsilon = b, A^\top y + s - Qx = c\\ &x_{\mathcal{B}(\epsilon^0)} \geq 0, x_{\mathcal{N}(\epsilon^0)\cup\mathcal{T}(\epsilon^0)} = 0, s_{\mathcal{N}(\epsilon^0)} \geq 0\\ &s_{\mathcal{B}(\epsilon^0)\cup\mathcal{T}(\epsilon^0)} = 0\}\\ u_{\epsilon^0} = \max_{\epsilon,x,y,s}\{&\epsilon : Ax - \Delta b\epsilon = b, A^\top y + s - Qx = c\\ &x_{\mathcal{B}(\epsilon^0)} \geq 0, x_{\mathcal{N}(\epsilon^0)\cup\mathcal{T}(\epsilon^0)} = 0, s_{\mathcal{N}(\epsilon^0)} \geq 0\\ &s_{\mathcal{B}(\epsilon^0)\cup\mathcal{T}(\epsilon^0)} = 0\}.\end{aligned}$$

Step 4: Recursively call

$$\begin{aligned}\Upsilon^0 &= \texttt{sub\_function}(A, b, \Delta b, c, Q, \underline{\epsilon}, l_{\epsilon^0})\\ \Upsilon^1 &= \texttt{sub\_function}(A, b, \Delta b, c, Q, u_{\epsilon^0}, \overline{\epsilon})\end{aligned}$$

and return $\Upsilon = \Upsilon^0 \cup \{l_{\epsilon^0}, u_{\epsilon^0}\} \cup \Upsilon^1$.

}

C. Calculating Impact Coefficients

Impact coefficients can be calculated by obtaining the system model output for all possible input variables and then solving a linear system of equations given below.

The realized values of LMPs according to the system model are determined by the combination of all factors; thus, they correspond to $p^*_{n,t}(1, 1, 1, 1; \epsilon_t)$. So, $\mu^n_{\mathrm{LMP}}(t) = \mu^n_t(1, 1, 1, 1)$ and $\sigma^n_{\mathrm{LMP}}(t) = \sigma^n_t(1, 1, 1, 1)$, where $\mu^n_{\mathrm{LMP}}(t)$ and $\sigma^n_{\mathrm{LMP}}(t)$ represent, respectively, the mean and standard deviation of LMPs at

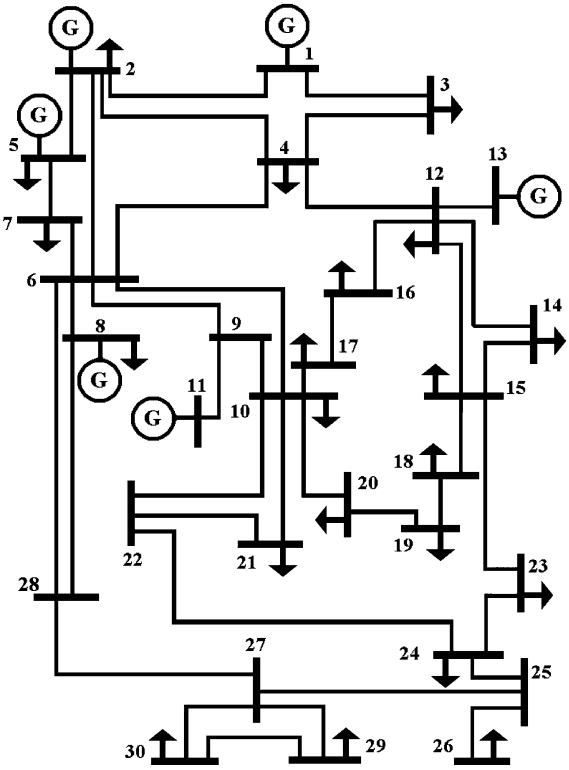

Fig. 2. IEEE 30-bus network example.

node n in hour t. Moreover, for all possible binary input variables $\mathbf{x} = \{x_1, x_2, x_3, x_4\}$, we have the following relation between the outputs and the impact coefficients:

$$\begin{aligned}\mu_t^n(x_1,x_2,x_3,x_4) =& \mu_0^n(t) + \mu_1^n(t)x_1 + \mu_2^n(t)x_2 + \mu_3^n(t)x_3 \\ &+ \mu_4^n(t)x_4 + \mu_{12}^n(t)x_1x_2 + \mu_{13}^n(t)x_1x_3 \\ &+ \mu_{14}^n(t)x_1x_4 + \mu_{23}^n(t)x_2x_3 \\ &+ \mu_{24}^n(t)x_2x_3 + \mu_{34}^n(t)x_3x_4 \\ &+ \mu_{123}^n(t)x_1x_2x_3 + \mu_{124}^n(t)x_1x_2x_4 \\ &+ \mu_{134}^n(t)x_1x_3x_4 + \mu_{234}^n(t)x_2x_3x_4 \\ &+ \mu_{1234}^n(t)x_1x_2x_3x_4, \\ \sigma_t^n(x_1,x_2,x_3,x_4) =& \sigma_0^n(t) + \sigma_1^n(t)x_1 + \sigma_2^n(t)x_2 + \sigma_3^n(t)x_3 \\ &+ \sigma_4^n(t)x_4 + \sigma_{12}^n(t)x_1x_2 + \sigma_{13}^n(t)x_1x_3 \\ &+ \sigma_{14}^n(t)x_1x_4 + \sigma_{23}^n(t)x_2x_3 \\ &+ \sigma_{24}^n(t)x_2x_3 + \sigma_{34}^n(t)x_3x_4 \\ &+ \sigma_{123}^n(t)x_1x_2x_3 + \sigma_{124}^n(t)x_1x_2x_4 \\ &+ \sigma_{134}^n(t)x_1x_3x_4 + \sigma_{234}^n(t)x_2x_3x_4 \\ &+ \sigma_{1234}^n(t)x_1x_2x_3x_4\end{aligned}$$

where $\mu_0^n(t), \mu_1^n(t), \ldots, \mu_{1234}^n(t)$ and $\sigma_0^n(t), \sigma_1^n(t), \ldots, \sigma_{1234}^n(t)$ are the impact coefficients at node n in hour t.

In matrix form, we have

$$Y = M \cdot \beta \tag{9}$$

where M, Y, and β are, respectively, a constant matrix, the system outputs, and the impact coefficients

$$M = \begin{bmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&1&0&0&1&0&0&0&0&0&0&0&0&0&0\\
1&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
1&1&0&1&0&0&1&0&0&0&0&0&0&0&0&0\\
1&0&1&1&0&0&0&0&1&0&0&0&0&0&0&0\\
1&1&1&1&0&1&1&0&1&0&0&1&0&0&0&0\\
1&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
1&1&0&0&1&0&0&1&0&0&0&0&0&0&0&0\\
1&0&1&0&1&0&0&0&0&1&0&0&0&0&0&0\\
1&1&1&0&1&1&0&1&0&1&0&0&1&0&0&0\\
1&0&0&1&1&0&0&0&0&0&1&0&0&0&0&0\\
1&1&0&1&1&0&1&1&0&0&1&0&0&1&0&0\\
1&0&1&1&1&0&0&0&1&1&1&0&0&0&1&0\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1
\end{bmatrix}$$

$$Y = \begin{bmatrix}
\mu_t^n(0,0,0,0)\\ \mu_t^n(1,0,0,0)\\ \mu_t^n(0,1,0,0)\\ \mu_t^n(1,1,0,0)\\
\mu_t^n(0,0,1,0)\\ \mu_t^n(1,0,1,0)\\ \mu_t^n(0,1,1,0)\\ \mu_t^n(1,1,1,0)\\
\mu_t^n(0,0,0,1)\\ \mu_t^n(1,0,0,1)\\ \mu_t^n(0,1,0,1)\\ \mu_t^n(1,1,0,1)\\
\mu_t^n(0,0,1,1)\\ \mu_t^n(1,0,1,1)\\ \mu_t^n(0,1,1,1)\\ \mu_t^n(1,1,1,1)
\end{bmatrix}, \quad \beta = \begin{bmatrix}
\mu_0^n(t)\\ \mu_1^n(t)\\ \mu_2^n(t)\\ \mu_3^n(t)\\ \mu_4^n(t)\\ \mu_{12}^n(t)\\ \mu_{13}^n(t)\\ \mu_{14}^n(t)\\ \mu_{23}^n(t)\\ \mu_{24}^n(t)\\ \mu_{34}^n(t)\\ \mu_{123}^n(t)\\ \mu_{124}^n(t)\\ \mu_{134}^n(t)\\ \mu_{234}^n(t)\\ \mu_{1234}^n(t)
\end{bmatrix}.$$

The matrix form for σ is similar. Therefore, if we obtain $\mu_t^n(x_1,x_2,x_3,x_4)$ and $\sigma_t^n(x_1,x_2,x_3,x_4)$ for all 16 possible inputs, then the impact coefficients in hour t can be calculated by solving the above linear system of (9).

The inverse of matrix M is shown in the equation at the bottom of the next page, which interprets how the impact coefficients are calculated using the outputs. For example, $\mu_1^n(t)$ is the difference of $\mu_t^n(1,0,0,0)$ and $\mu_t^n(0,0,0,0)$; $\mu_{12}^n(t)$ is the difference of $\mu_t^n(0,0,0,0) + \mu_t^n(1,1,0,0)$ and $\mu_t^n(0,1,0,0) + \mu_t^n(1,0,0,0)$.

IV. Numerical Example

A. Test Data

We use the IEEE 30-bus network as an example to demonstrate our approach. As is shown in Fig. 2, a supply node has a "G" in a circle representing a generator, and a demand node has an arrow. In this example, $\mathcal{D} = \{2,3,4,5,7,8,10, 12,14,15,16,17,18,19,20,21,23,24,26,29,30\}$, and $\mathcal{S} =$

TABLE II
NODE DATA OF THE 30-BUS EXAMPLE

n	d_n	a_n	α_n	b_n	$\overline{Q_n}$
1	–	0	14.00	1.00	160
2	15.19	0	12.10	0.88	64
3	1.68	–	–	–	–
4	5.32	–	–	–	–
5	65.94	0	8.28	0.50	40
6	–	–	–	–	–
7	15.96	–	–	–	–
8	21	0	16.58	1.63	28
9	–	–	–	–	–
10	4.06	–	–	–	–
11	–	0	15.41	1.50	24
12	7.84	–	–	–	–
13	–	0	15.39	1.50	32
14	4.34	–	–	–	–
15	5.74	–	–	–	–
16	2.45	–	–	–	–
17	6.3	–	–	–	–
18	2.24	–	–	–	–
19	6.65	–	–	–	–
20	1.54	–	–	–	–
21	12.25	–	–	–	–
22	–	–	–	–	–
23	2.24	–	–	–	–
24	6.09	–	–	–	–
25	–	–	–	–	–
26	2.45	–	–	–	–
27	–	–	–	–	–
28	–	–	–	–	–
29	1.68	–	–	–	–
30	7.42	–	–	–	–

TABLE III
TRANSMISSION LINE DATA OF THE 30-BUS EXAMPLE

line	resistance R (Ω)	reactance X (Ω)	thermal limit (MW)
1-2	0.0192	0.0575	78
1-3	0.0452	0.1852	78
2-4	0.0570	0.1737	39
3-4	0.0132	0.0379	78
2-5	0.0472	0.1983	78
2-6	0.0581	0.1763	39
4-6	0.0119	0.0414	54
5-7	0.0460	0.1160	42
6-7	0.0267	0.0820	78
6-8	0.0120	0.0420	19.2
6-9	0	0.2080	39
6-10	0	0.5560	19.2
9-11	0	0.2080	39
9-10	0	0.1100	39
4-12	0	0.2560	39
12-13	0	0.1400	39
12-14	0.1231	0.2559	19.2
12-15	0.0662	0.1304	19.2
12-16	0.0945	0.1987	19.2
14-15	0.2210	0.1997	9.6
16-17	0.0824	0.1932	9.6
15-18	0.1070	0.2185	9.6
18-19	0.0639	0.1292	9.6
19-20	0.0340	0.0680	19.2
10-20	0.0936	0.2090	19.2
10-17	0.0324	0.0845	19.2
10-21	0.0348	0.0749	19.2
10-22	0.0727	0.1499	19.2
21-22	0.0116	0.0236	19.2
15-23	0.1000	0.2020	9.6
22-24	0.1150	0.1790	9.6
23-24	0.1320	0.2700	9.6
24-25	0.1885	0.3292	9.6
25-26	0.2544	0.3800	9.6
25-27	0.1093	0.2087	9.6
28-27	0	0.3960	39
27-29	0.2198	0.4153	9.6
27-30	0.3202	0.6027	9.6
29-30	0.2399	0.4533	9.6
8-28	0.0636	0.2000	19.2
6-28	0.0169	0.0599	19.2

$\{1, 2, 5, 8, 11, 13\}$. Node and transmission line data are given in Tables II and III, respectively. Most of these data are adopted from [15] and [19]: b_n's, T_l's, and $\overline{Q_n}$'s are set to be, respectively, 50%, 60%, and 80% of the values in [19]; α_n's are set to be 50% of the equilibrium values in [15]. For all $n \in \mathcal{D}$, the average nominal load over time horizon is set to be 70% of d_n in [19]; the relative load chronological changes, shown in Fig. 3, are estimated using the load data of PJM-E [20]; the demand uncertainty ϵ_t in each hour is assumed to have a truncated normal distribution also estimated using the load data of PJM-E. Let $f_N(\cdot)$ be the pdf of a normal distribution, and

$$M^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 1 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 & 0 & 0 & -1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1
\end{bmatrix}$$

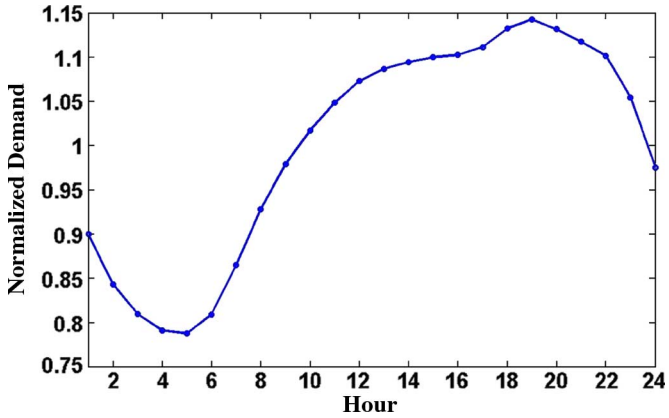

Fig. 3. Normalized chronological load change.

TABLE IV
SELECTED IMPACT COEFFICIENTS FOR $\mu^n_{\rm LMP}$

n	$\mu^n_{\rm LMP}$	μ^n_0	μ^n_1	μ^n_2	μ^n_3	μ^n_4
1	61.80	39.61	2.09	0	3.57	14.12
2	61.81	39.61	2.09	0	3.57	14.12
5	61.85	39.61	2.09	0	3.57	14.12
8	61.89	39.61	2.09	0	3.57	14.12
11	62.23	39.61	2.09	0	3.57	14.12
13	60.80	39.61	2.09	0	3.57	14.12
15	64.99	39.61	2.09	0	3.57	14.12
18	64.08	39.61	2.09	0	3.57	14.12
23	64.18	39.61	2.09	0	3.57	14.12
30	62.36	39.61	2.09	0	3.57	14.12
average	62.43	39.61	2.09	0	3.57	14.12
n	μ^n_{13}	μ^n_{14}	μ^n_{34}	μ^n_{123}	μ^n_{134}	μ^n_{1234}
1	2.66	–0.11	–0.32	–0.01	0.19	0.00
2	2.66	–0.11	–0.32	0.02	0.19	–0.01
5	2.66	–0.11	–0.32	0.10	0.19	–0.06
8	2.66	–0.11	–0.32	0.21	0.19	–0.12
11	2.66	–0.11	–0.32	0.97	0.19	–0.55
13	2.66	–0.11	–0.32	–2.26	0.19	1.26
15	2.66	–0.11	–0.32	7.18	0.19	–4.00
18	2.66	–0.11	–0.32	5.14	0.19	–2.86
23	2.66	–0.11	–0.32	5.37	0.19	–2.99
30	2.66	–0.11	–0.32	1.26	0.19	–0.71
average	2.66	–0.11	–0.32	1.40	0.19	–0.79

then, the pdf of a truncated normal distribution $f_T(\cdot)$ with the same mean and standard deviation within $[\underline{x}, \overline{x}]$ is

$$f_T(x) = \begin{cases} f_N(x) / \int_{\underline{x}}^{\overline{x}} f_N(y)dy, & \underline{x} \le x \le \overline{x} \\ 0, & \text{otherwise.} \end{cases}$$

B. Time-Averaged Impact Coefficients

We obtain the impact coefficients for both mean and standard deviation of LMPs at all nodes $n \in \mathcal{D} \cup \mathcal{S}$ in hour $t = 1, \ldots, 24$. Tables IV and V show the time-averaged impact coefficients for $\mu^n_{\rm LMP}$ and $\sigma^n_{\rm LMP}$, respectively. Here

$$\mu^n_{\rm LMP} = \frac{1}{24} \sum_{t=1}^{24} \mu^n_{\rm LMP}(t).$$

Time averages are also taken in the same manner for other impact coefficients. For the purpose of conserving more space, only some significant impact coefficients for selected nodes are given. The "average" rows are the average values over all nodes

TABLE V
SELECTED IMPACT COEFFICIENTS FOR $\sigma^n_{\rm LMP}$

n	$\sigma^n_{\rm LMP}$	σ^n_1	σ^n_{13}	σ^n_{14}	σ^n_{123}	σ^n_{134}	σ^n_{1234}
1	5.62	3.24	2.34	–0.11	0.00	0.16	0.00
2	5.63	3.24	2.34	–0.11	0.02	0.16	–0.01
5	5.66	3.24	2.34	–0.11	0.07	0.16	–0.03
8	5.69	3.24	2.34	–0.11	0.14	0.16	–0.07
11	5.94	3.24	2.34	–0.11	0.63	0.16	–0.31
13	5.01	3.24	2.34	–0.11	–1.36	0.16	0.75
15	8.09	3.24	2.34	–0.11	4.58	0.16	–2.11
18	7.37	3.24	2.34	–0.11	3.28	0.16	–1.53
23	7.45	3.24	2.34	–0.11	3.43	0.16	–1.60
30	6.04	3.24	2.34	–0.11	0.81	0.16	–0.40
average	6.11	3.24	2.34	–0.11	0.91	0.16	–0.42

in $n \in \mathcal{D} \cup \mathcal{S}$. We have the following observations and interpretations based on these results.

1) Had none of the four factors existed, the LMPs would have been constantly and universally \$39.61/MWh, compared to the realized mean of \$62.43/MWh with a \$6.11/MWh standard deviation on average.
2) As the sole contributions of the individual factors, load uncertainty, capacity reserve, and market power raise the mean of LMPs by $\mu^n_1 = \$2.09/\text{MWh}$, $\mu^n_3 = \$3.57/\text{MWh}$, and $\mu^n_4 = \$14.12/\text{MWh}$, respectively.
3) The factor thermal limit does not increase LMPs by itself ($\mu^n_2 = \$0/\text{MWh}$), nor does the interaction between load uncertainty and thermal limit ($\mu^n_{12} = \$0/\text{MWh}$). This is because in this particular example, the transmission capacity is sufficient when 90% criterion is inactive. This result, however, may not necessarily hold in general, and μ^n_2 and/or μ^n_{12} could become non-zeros for other network settings.
4) Price difference between nodes is an indication of congestion. Only two impact coefficients μ^n_{123} and μ^n_{1234} in Table IV take different values over different nodes, which means that congestion is not caused by a single factor; rather, it is a result of the interaction among three or four factors. For the same reason as explained above, there may exist other combinations of factors that also contribute to the congestion, but the significant source of congestion is still believed to be μ^n_{123} and μ^n_{1234}. However, those columns where factor 2 does not appear (μ^n_0, μ^n_1, μ^n_{34}, etc.) can be proven to be constant across nodes regardless of system parameters, because the only cause of price difference in a dc lossless model, thermal limit (factor 2), is set to be inactive in the computation of these columns.
5) It is also interesting to observe that μ^n_{1234}'s have smaller magnitudes with opposite signs than those of μ^n_{123}'s. This means that, given the existence of the first three factors, the incremental effect of market power mitigates congestion. One possible interpretation of this phenomenon is that market power reduces the relative differences among supply functions and thus diminishes the preference for less-expensive generators. Comparing the supply function parameters α_n's and b_n's in Table II, we find that they have a positive correlation, but α_n's are less spread out than b_n's

$$b_5 : b_2 : b_1 : b_{13} : b_{11} : b_8 = 1 : 1.75 : 2 : 3 : 3 : 3.25$$
$$\alpha_5 : \alpha_2 : \alpha_1 : \alpha_{13} : \alpha_{11} : \alpha_8 = 1 : 1.46 : 1.69 : 1.86 : 1.86 : 2.00.$$

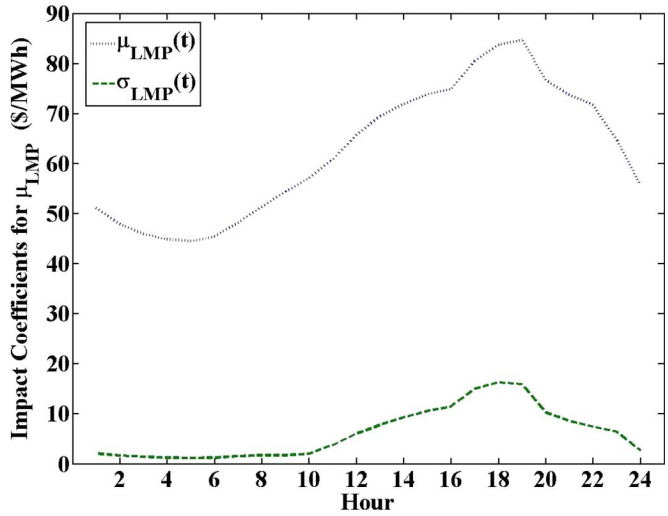

Fig. 4. Time variation of $\mu_{\mathrm{LMP}}(t)$ and $\sigma_{\mathrm{LMP}}(t)$.

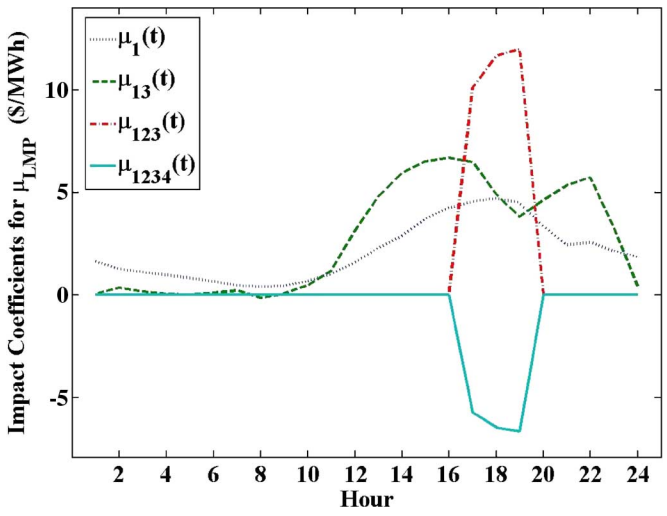

Fig. 5.
Time variation of impact coefficients for $\mu_{\mathrm{LMP}}(t)$.

Without market power, the ratio of supply functions of a more expensive generator i and a less expensive one j is

$$\frac{a_i q + b_i q^2/2}{a_j q + b_j q^2/2} = \frac{b_i}{b_j} > 1.$$

When market power becomes active, α_n's substitute zero-valued a_n's, and this ratio becomes

$$\frac{\alpha_i q + \frac{1}{2} b_i q^2}{\alpha_j q + \frac{1}{2} b_j q^2} < \frac{\left(\alpha_j \frac{b_i}{b_j}\right) q + \frac{1}{2} b_i q^2}{\alpha_j q + \frac{1}{2} b_j q^2} = \frac{b_i}{b_j}$$

which means that the relative differences among supply functions shrink. This phenomenon has also been observed and discussed in [21]. However, we can only conclude that market power could mitigate congestion in certain cases under certain assumptions (e.g., how market power is exercised) but not necessarily so in all circumstances.

6) Load uncertainty is the single primary and decisive source of LMP volatility, contributing an average standard deviation of $\sigma_1^n = \$3.24/\mathrm{MWh}$ to the realized \$6.11/MWh in total. The absence of sole contributions of other factors ($\sigma_2^n = \sigma_3^n = \sigma_4^n = \$0/\mathrm{MWh}$) is due to their assumed deterministic characteristics. The interactions between these factors and load uncertainty, however, have significant contribution to the LMP volatility. It is also indicated in Table V that congestion is a result of the interaction among factors.

C. Time Variation of Impact Coefficients

We show in Fig. 4 how the node-averaged mean and standard deviation of LMPs vary from hour to hour. In Fig. 4

$$\mu_{\mathrm{LMP}}(t) = \frac{1}{|\mathcal{D} \cup \mathcal{S}|} \sum_{n \in \mathcal{D} \cup \mathcal{S}} \mu_{\mathrm{LMP}}^n(t).$$

Node-averages are also taken in the same manner for the impact coefficients, which are shown in Figs. 5 and 6. Only some of the impact coefficients are plotted; the omitted ones have little or no variation over time.

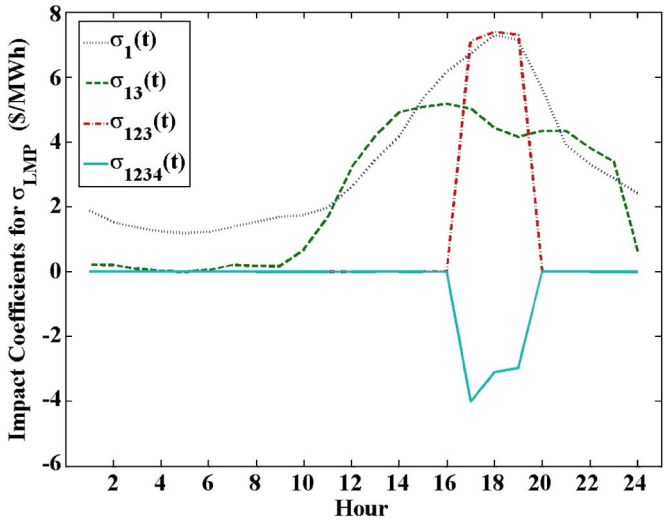

Fig. 6. Time variation of impact coefficients for $\sigma_{\mathrm{LMP}}(t)$.

We have the following observations.

1) Coefficients $\mu_{\mathrm{LMP}}(t)$ and $\mu_1(t)$ follow a similar pattern of time variation with $d_{n,t}$. Recall that $\mu_1(t)$ represents the sole contribution of factor 1 at node n in hour t, and it is the difference between $\mu_t(x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0)$ and $\mu_t(x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0)$, which represent the two situations with demand at node n in hour t being $d_{n,t}(1 + \epsilon_t)$ and $d_{n,t}$.
2) Coefficient $\mu_{13}(t)$, interaction between factors 1 and 3, becomes significant when demand is high.
3) We can see from $\mu_{123}(t)$ and $\mu_{1234}(t)$ that congestion occurs during peak hours and that market power mitigates the congestion.
4) The observations on the impact coefficients for σ_{LMP} are similar.

V. Use of System Model in Evaluating Expansion Plans

The system model approach described in Section III can be used to perform sensitivity analysis for the current system, or to analyze the effects of other factors or system upgrading decisions. As an illustration, we use the network example in Section IV to analyze the effects of generator and transmission line capacity expansion plans.

Suppose investment decisions are to be made to introduce new generators at nodes 15, 18, and 23, where highest means and largest standard deviations of LMPs are observed. The supply function parameters of the new generators are all assumed to take the average values of the previously existing ones, i.e., $a_n = 0$, $\alpha_n = 13.63$, $b_n = 1.17$, $\overline{Q_n} = 52.2$, $\forall n = 15, 18, 23$. Transmission line 12–15 is also planned to be expanded by 15% of its current capacity. The questions are: (1) how would the LMPs be affected, and (2) how effective would the capacity expansion of each component (generator or transmission line) be?

Keeping all of the previous factors (load uncertainty, thermal limit, capacity reserve, and market power) active, we define another four factors, numbered 5, 6, 7, and 8, as the introduction of generators 15, 18, and 23, and capacity expansions of transmission line 12–15, respectively.

The statistical models for this problem become

$$\begin{aligned}\tilde{\mu}^n_{\rm LMP}(t) =& \mu^n_{\rm LMP}(t) + \mu^n_5(t) + \mu^n_6(t) + \mu^n_7(t) + \mu^n_8(t)\\ &+ \mu^n_{56}(t) + \mu^n_{57}(t) + \mu^n_{58}(t) + \mu^n_{67}(t) + \mu^n_{68}(t)\\ &+ \mu^n_{78}(t) + \mu^n_{567}(t) + \mu^n_{568}(t) + \mu^n_{578}(t)\\ &+ \mu^n_{678}(t) + \mu^n_{5678}(t)\\ \tilde{\sigma}^n_{\rm LMP}(t) =& \sigma^n_{\rm LMP}(t) + \sigma^n_5(t) + \sigma^n_6(t) + \sigma^n_7(t) + \sigma^n_8(t)\\ &+ \sigma^n_{56}(t) + \sigma^n_{57}(t) + \sigma^n_{58}(t) + \sigma^n_{67}(t) + \sigma^n_{68}(t)\\ &+ \sigma^n_{78}(t) + \sigma^n_{567}(t) + \sigma^n_{568}(t) + \sigma^n_{578}(t)\\ &+ \sigma^n_{678}(t) + \sigma^n_{5678}(t)\end{aligned}$$

where $\tilde{\mu}^n_{\rm LMP}(t)$ and $\tilde{\sigma}^n_{\rm LMP}(t)$ are, respectively, the forecasted mean and standard deviation of LMP at node n as a result of these additional factors; as has been defined before, $\mu^n_{\rm LMP}(t)$ and $\sigma^n_{\rm LMP}(t)$ are, respectively, the realized mean and standard deviation before the introduction of new factors.

We obtain the impact coefficients at all nodes $n \in \mathcal{D} \cup \mathcal{S}$ in each hour $t = 1, \ldots, 24$. Tables VI and VII show the time-averaged impact coefficients for $\tilde{\mu}^n_{\rm LMP}$ and $\tilde{\sigma}^n_{\rm LMP}$, respectively. As a result of the combined effect of these four factors, the mean and standard deviation of LMPs are reduced on average from \$62.43/MWh and \$6.11/MWh to \$49.18/MWh and \$3.24/MWh, respectively. The individual contributions of factors 5, 6, 7, and 8 to $\tilde{\mu}^n_{\rm LMP}$ are, respectively, −\$6.71/MWh, −\$7.61/MWh, −\$8.84/MWh, and −\$0.62/MWh on average; and their individual contributions to $\tilde{\sigma}^n_{\rm LMP}$ are, respectively, −\$0.49/MWh, −\$1.82/MWh, −\$2.19/MWh, and −\$0.49/MWh on average.

Tables VI and VII can also provide information for any subset of factors 5, 6, 7, and 8. For example, to answer the question of "what is the incremental value of expanding the capacity of transmission line 12–15 by 15% after new generators have been introduced at nodes 15, 18 and 23?", we select those columns in (unabbreviated) Table VI that contain factor 8: $\{\mu_8, \mu_{58}, \mu_{68}, \mu_{78}, \mu_{568}, \mu_{578}, \mu_{678}, \mu_{5678}\}$. The summation of these columns is zero. The same result is observed for Table VII. It indicates that the incremental value of expanding the capacity of transmission line 12–15 beyond introducing new generators is zero. This information is useful for decision makers to avoid redundant investments.

By redefining the factors as unit increments of certain system parameters (e.g., b_n's as fuel prices increase or α_n's as more severe market power), sensitivity analysis for the current system can also be performed in a similar way as illustrated in this section.

TABLE VI
Selected Impact Coefficients for $\tilde{\mu}^n_{\rm LMP}$

n	$\tilde{\mu}^n_{\rm LMP}$	$\mu^n_{\rm LMP}$	μ^n_5	μ^n_6	μ^n_7	μ^n_8
1	50.10	61.80	–8.82	–6.79	–7.44	0.01
2	50.24	61.81	–8.74	–6.73	–7.38	0.00
5	50.60	61.85	–8.53	–6.58	–7.22	–0.04
8	51.08	61.90	–8.32	–6.44	–6.95	–0.09
11	54.45	62.23	–5.18	–4.13	–6.33	–0.43
13	40.20	60.80	–14.62	–10.93	–11.83	1.01
15	31.44	64.99	–25.21	–19.97	–20.48	–3.18
18	22.36	64.08	28.72	–33.19	–15.35	–2.27
23	23.53	64.18	–18.70	–15.10	–31.94	–2.38
30	55.70	62.36	–9.23	–7.27	–2.21	–0.56
average	49.18	62.43	–6.71	–7.61	–8.84	–0.62
n	μ^n_{56}	μ^n_{57}	μ^n_{58}	μ^n_{67}	μ^n_{567}	μ^n_{5678}
1	5.53	6.37	–0.01	3.91	–4.46	–0.01
2	5.50	6.34	0.00	3.90	–4.47	0.00
5	5.42	6.28	0.04	3.85	–4.47	0.04
8	5.41	6.24	0.09	3.79	–4.55	0.09
11	3.01	5.09	0.43	3.39	–3.63	0.43
13	7.79	8.08	–1.01	5.14	–4.24	–1.01
15	15.41	14.86	3.18	11.29	–9.44	3.18
18	–36.82	–9.29	2.27	12.26	11.95	2.27
23	33.38	10.59	2.38	12.40	31.29	2.38
30	9.43	7.18	0.56	3.17	–7.74	0.56
average	4.26	5.73	0.62	4.75	–4.83	0.62

TABLE VII
Selected Impact Coefficients for $\tilde{\sigma}^n_{\rm LMP}$

n	$\tilde{\sigma}^n_{\rm LMP}$	$\sigma^n_{\rm LMP}$	σ^n_5	σ^n_6	σ^n_7	σ^n_8
1	3.28	5.62	–1.24	–1.33	–1.58	0.00
2	3.30	5.63	–1.22	–1.32	–1.58	–0.01
5	3.35	5.66	–1.14	–1.31	–1.56	–0.03
8	3.42	5.69	–1.07	–1.31	–1.54	–0.07
11	3.88	5.94	0.08	–1.04	–1.59	–0.32
13	1.89	5.01	–3.38	–1.73	–2.06	0.61
15	0.66	8.09	–6.53	–5.75	–6.04	–2.47
18	0.54	7.37	13.23	–6.94	–4.47	–1.75
23	0.53	7.45	–5.67	–4.32	–7.03	–1.83
30	4.09	6.04	–1.59	–1.73	–0.83	–0.41
average	3.24	6.11	–0.49	–1.82	–2.19	–0.49
n	σ^n_{56}	σ^n_{57}	σ^n_{58}	σ^n_{67}	σ^n_{567}	σ^n_{5678}
1	0.99	1.35	0.00	0.79	–1.31	0.00
2	0.98	1.34	0.01	0.79	–1.32	0.01
5	0.96	1.32	0.03	0.79	–1.36	0.03
8	1.00	1.30	0.07	0.79	–1.44	0.07
11	–0.09	1.21	0.32	0.80	–1.43	0.32
13	1.57	1.90	–0.61	0.87	–0.31	–0.61
15	5.45	5.99	2.47	4.56	–5.12	2.47
18	–13.33	1.65	1.75	4.39	–1.37	1.75
23	18.43	5.77	1.83	4.22	–18.32	1.83
30	3.66	0.90	0.41	0.83	–3.19	0.41
average	0.89	1.75	0.49	1.32	–2.33	0.49

VI. Conclusion

This paper builds a system model to decompose the effects of influential factors on locational marginal prices. Four factors (load uncertainty, thermal limit, capacity reserve, and market power) are considered, and the impact coefficients are calculated to estimate the contribution of each single factor and their interactions to the mean and standard deviation of LMPs at each node.

An IEEE 30-bus network is used as an example to demonstrate this approach. The system model approach can also be used to perform sensitivity analysis or to evaluate the effectiveness of investment plans, e.g, introducing new generators and/or expanding capacities of transmission lines.

The main contributions of this paper include the following.

1) The system model enables one to answer "what if" questions, which are generally hard to answer using historical data.
2) Piecewise linear property of LMPs as functions of demand variation has been explored. For a given continuous probability distribution of demand uncertainty and input variables, the mean and standard deviation of LMPs can be obtained exactly and efficiently using integration, which is equivalent to infinitely many simulation samples if Monte Carlo simulation were to be used instead.
3) Impact coefficients provide insights on the composition of LMP probability distribution, in terms of mean and standard deviation. They can also inform and assist power system evaluation and investment decision making. The techniques in this paper can also be used for a more complete analysis of the LMP probability distribution (e.g., on higher moments).

It is worth mentioning that our observations and analyses are based on the consideration of only four factors with simplifying assumptions and are only derived for some specified system parameters. Further research should (1) consider other factors that affect LMP probability distributions, e.g., fuel prices fluctuation, generator, or transmission line outage; (2) relax the assumption of perfect load correlation among nodes (according to the load data from PJM [20], the correlation between PJM-E and PJM-W is 0.8632); (3) relax the assumption that supply functions stay the same over the entire time horizon, and study the dynamic gaming behavior of market power exercise under different scenarios and inputs; and (4) compare the analyses from the system model to real-world observations.

Acknowledgment

The authors would like to thank the editors and four anonymous referees for their insightful and constructive comments.

References

[1] F. Schweppe, M. Caramanis, R. Tabors, and R. Bohn, *Spot Pricing of Electricity*. Boston, MA: Kluwer, 1998.

[2] J. Valenzuela and M. Mazumdar, "A probability model for the electricity price duration curve under an oligopoly market," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1250–1256, Aug. 2005.

[3] A. Conejo, E. Castillo, R. Mínguez, and F. Milano, "Locational marginal price sensitivities," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 2026–2033, Nov. 2005.

[4] B. Lesieutre, H. Oh, R. Thomas, and V. Donde, "Identification of market power in large-scale electric energy markets," in *Proc. 39th Hawaii Int. Conf. System Sciences*, 2006, vol. 10, p. 240b.

[5] M. Burger, B. Klar, A. Müller, and G. Schindlmayr, "A spot market model for pricing derivatives in electricity markets," *Quant. Fin.*, vol. 4, pp. 109–122, 2004.

[6] F. Longstaff and A. Wang, "Electricity forward prices: A high-frequency empirical analysis," *J. Fin.*, vol. 59, no. 4, pp. 1877–1900, 2004.

[7] J. Popova, "Spatial pattern in modeling electricity prices: Evidence from the PJM market," in *Proc. 24th USAEE and IAEE North American Conf.*, Washington, DC, Jul. 8–10, 2004.

[8] W. Hogan, "Markets in real electric networks require reactive prices," *Energy J.*, vol. 14, no. 3, pp. 171–200, 1993.

[9] T. Overbye, X. Cheng, and Y. Sun, "A comparison of the AC and DC power flow models for LMP calculations," in *Proc. 37th Hawaii Int. Conf. System Sciences*, 2004.

[10] Y. Chen and B. Hobbs, "An oligopolistic power market model with tradable NO_x permits," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 119–129, Feb. 2005.

[11] L. Wang, M. Mazumdar, M. Bailey, and J. Valenzuela, "Oligopoly models for market price of electricity under demand uncertainty and unit reliability," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1309–1321, 2007.

[12] F. Bouffard, F. Galiana, and A. Conejo, "Market-clearing with stochastic security—Part I: Formulation," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1818–1826, Nov. 2005.

[13] F. Bouffard, F. Galiana, and A. Conejo, "Market-clearing with stochastic security—Part II: Case studies," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1827–1835, Nov. 2005.

[14] K. Dragoon and V. Dvortsov, "Z-method for power system resource adequacy applications," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 982–988, May 2006.

[15] B. Hobbs, C. Metzler, and J. Pang, "Strategic gaming analysis for electric power system: An MPEC approach," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 638–645, May 2000.

[16] A. Berkelaar, B. Jansen, K. Roos, and T. Terlaky, Sensitivity Analysis in (Degenerate) Quadratic Programming, Econometric Institute, Erasmus University Rotterdam, The Netherlands, Tech. Rep., 1996.

[17] O. Güler and Y. Ye, "Convergence behavior of interior-point algorithms," *Math. Program.*, vol. 60, pp. 215–228, 1993.

[18] L. McLinden, "An analogue of Moreau's approximation theorem, with application to the nonlinear complementary problem," *Pacific J. Math.*, vol. 88, pp. 101–161, 1980.

[19] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 745–751, May 1974.

[20] [Online]. Available: http://www.pjm.com.

[21] T. Li and M. Shahidehpour, "Market power analysis in electricity markets using supply function equilibrium model," *IMA J. Manage. Math.*, vol. 15, pp. 339–354, 2004.

Lizhi Wang received the B.E. degree in automation and the B.S. degree in management science from the University of Science and Technology of China, Heifei, China, in 2003 and the Ph.D. degree in industrial engineering from the University of Pittsburgh, Pittsburgh, PA, in 2007.

He is currently an Assistant Professor in the Department of Industrial and Manufacturing Systems Engineering and the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA.

Mainak Mazumdar received the Ph.D. degree in industrial engineering from Cornell University, Ithaca, NY, in 1966.

He worked as a Research Scientist at the Westinghouse R&D Center during the period 1966–1981. Since 1981, he has been a Professor in the Department of Industrial Engineering at the University of Pittsburgh, Pittsburgh, PA.