Learning the LMP-Load Coupling From Data: A Support Vector Machine Based Approach

Xinbo Geng, *Student Member, IEEE*, and Le Xie, *Senior Member, IEEE*

Abstract—This paper investigates the fundamental coupling between loads and locational marginal prices (LMPs) in security-constrained economic dispatch (SCED). Theoretical analysis based on multi-parametric programming theory points out the unique one-to-one mapping between load and LMP vectors. Such one-to-one mapping is depicted by the concept of system pattern region (SPR) and identifying SPRs is the key to understanding the LMP-load coupling. Built upon the characteristics of SPRs, the SPR identification problem is modeled as a classification problem from a market participant's viewpoint, and a Support Vector Machine based data-driven approach is proposed. It is shown that even without the knowledge of system topology and parameters, the SPRs can be estimated by learning from historical load and price data. Visualization and illustration of the proposed data-driven approach are performed on a 3-bus system as well as the IEEE 118-bus system.

Index Terms—Locational marginal prices (LMPs), security constrained economic dispatch (SCED), support vector machine (SVM), system pattern region.

I. INTRODUCTION

A FUNDAMENTAL issue with electricity market operation is to understand the impact of operating conditions (e.g., load levels at each bus) on the locational marginal prices (LMPs). This paper examines this key issue of the relationship between nodal load levels and LMPs. This issue is further compounded by the increasing levels of demand response and variable resources in the grid.

In the power systems literature, reference [1] is among the pioneering works that uses perturbation techniques to compute the sensitivities of the dual variables in SCED (e.g., LMPs) with respect to parameters (e.g., the nodal load levels). This sensitivity calculation method is widely used in subsequent researches. However, this approach is valid only for small changes and the marginal generator stays the same. Reference [2] observed the "step changes" of LMPs with respect to increasing system load level and discovered that new binding constraints (transmission or generation) are the reason of the "step changes". This

is followed by further analysis on identifying the critical load levels (CLLs) that trigger such step changes of LMPs [3]–[5]. This line of work assumes that the system load change is distributed to each bus proportional to the base case load, which, in many instances, do not necessarily represent the real-world situations. Reference [6] analyzed this problem using quadratic-linear programming and the concepts of *system patterns* and *system pattern regions (SPRs)* were first introduced. The SPRs depict the relationship between loads and LMPs in the whole load space, which is not confined in a small neighborhood of an operating point or constrained by a specific load distribution pattern. This paper is inspired by [6] but focuses on the case of piecewise linear generation costs, instead of the quadratic cost case in [6]. The reason that we study the piecewise linear cost case is that piecewise linear cost curves are often quite representative of the market practice in the real world. In addition, some new theoretical results based on piecewise linear cost curves are derived, and are generalizable towards quadratic cost cases.

Characterizing the SPRs would provide important insights to both system operators and market participants. Reference [7] advances the theory of SPR from system operator's perspective where the knowledge of system topology and parameters is available. For market participants, such knowledge is not necessarily available. Our previous work [8] examines the issue from market participant's viewpoint and applies the geometric features of SPRs to identify them.

This paper significantly advances our previous work by (1) completing the theoretical characterization of SPRs as a function of nodal load levels; (2) proposing a computational algorithm to identify SPRs using historical data; (3) introducing the posterior probabilities of SPRs with the presence of uncertain system parameters such as transmission limits; and (4) extending the algorithm to consider practical factors such as partial load information and loss component of LMPs.

The rest of the paper is organized as follows. Section II provides the analysis of LMP-load coupling in SCED problem from the viewpoint of MLP theory, with an illustrative example. Section III illustrates the changes of SPRs given changes of system parameters such as transmission limits. Based on the theoretical analysis, a data-driven algorithm for market participants to identify SPRs is described in Section IV. Section V illustrates the performance of the algorithm on the IEEE 118-bus system. Section VI explores the impact of nodal load information, and Section VII provides critical assessment of the

proposed method. Concluding remarks and future works are presented in Section VIII.

II. THEORETICAL ANALYSIS

A. Notations

The notations of this paper are summarized below: mathematical symbols in hollowed-out shapes (e.g., $\mathbb{R}$) represent spaces and symbols in Calligra font (e.g., $\mathcal{S}_\pi$) stand for sets. The superscript "$*$" indicates the variable is optimal, " ^ " denotes estimated values (e.g., $\hat{\lambda}$). Variables with " ¯ " are expectations or average values (e.g., $\bar{\lambda}$). "$\intercal$" denotes the transpose of a vector or matrix (e.g., $\mathbf{1}_n^\intercal$). The subscript "i" represents the ith element of the vector (e.g., P_{G_i}), and the superscript "(i)" represents the ith element in a set (e.g., $P_D^{(i)}$). The vector of $n \times 1$ ones, matrix of $m \times n$ zeros and the $n \times n$ identity matrix are denoted by $\mathbf{1}_n$ and $\mathbf{0}_{m \times n}$ and $\mathbf{I}_n$ respectively.

B. Security Constrained Economic Dispatch (SCED)

In real-time energy market operations, the LMPs are the results from the SCED, which is formulated as follows:

$$\min_{P_G^{(k)}} \quad \sum_{i=1}^{n_b} c_i(P_{G_i}^{(k)}) \tag{1a}$$

$$\text{s.t.} \quad \sum_{i=1}^{n_b} P_{G_i}^{(k)} = \sum_{j=1}^{n_b} P_{D_j}^{(k)} \qquad : \lambda_1 \tag{1b}$$

$$-F^+ \le H(P_G^{(k)} - P_D^{(k)}) \le F^+ \qquad : \mu^+, \mu^- \tag{1c}$$

$$P_G^- \le P_G^{(k)} \le P_G^+ \qquad : \eta^+, \eta^- \tag{1d}$$

where $P_G^{(k)}$ is the generation vector at time k, and $P_D^{(k)}$ is the load vector at time k. We assume there are both generation and load at each bus. Let n_b denote the number of buses and n_l denote the number of transmission lines, then $P_G^{(k)}, P_D^{(k)} \in \mathbb{R}^{n_b}$. $H \in \mathbb{R}^{n_l \times n_b}$ is the shift factor matrix.

This formulation considers each snapshot independently, therefore it is called *static SCED* in this paper. For simplicity, we write $P_G^{(k)}$ and $P_D^{(k)}$ as P_G and P_D when discussing the static SCED.

The objective of SCED is to minimize the total generation cost and satisfy the transmission and generation capacity constraints while keeping the real-time balance between supply and demand. The generation cost function $c_i(P_{G_i}^{(k)})$ of generator i is increasing and convex, and it is usually regarded as a quadratic function or approximated by a piecewise linear function. To better reflect the current practice in electricity markets, this paper studies the SCED problem with piecewise linear generator bidding functions. And for the consideration of simplicity, the simplest form, i.e., $\sum_{i=1}^{n_b} c_i(P_{G_i}) = c^\intercal P_G$ is being considered in this paper.

A fundamental concept in electricity markets is the *LMP*. The LMP λ_i at bus i is defined as the change of total system cost if the demand at node i is increased by 1 unit [9]. According to [10], the LMP vector λ can be calculated by the following equation:

$$\lambda = \lambda_1 \mathbf{1}_{n_b} + H^\intercal(\mu^+ - \mu^-) \tag{2}$$

We start with the simplest case of static SCED. More elaborated SCED formulations are in Section V-D. Since line losses are not explicitly modeled in the SCED formulation, the LMPs in this paper do not contain the loss components. Further discussions on the loss component are in Section VII-D.

C. SCED Analysis via Multiparametric Linear Programming (MLP)

In real-world market operations, the parameters associated with the SCED above are typically time-varying. Therefore, it is essential to understand the effects of parameters on the optimality of the problem. Multiparametric Programming problem aims at exploring the characteristics of an optimization problem which depending on a *vector of parameters* [11]. MLP theory, which is the foundation of this paper, pays special attention to Linear Programming problems.

The objective of this paper is to understand the impact of parameters (i.e., load levels, line capacities, etc.) on the outcome of SCED (namely, the prices). We pose the problem in view of MLP, and analyze the theoretical properties.

In reality, LMP vectors depend upon a number of factors, including: (1) the loads in the system; (2) line flow limits; (3) ramp constraints; (4) generation offer prices; (5) topology of the system; (6) unit commitment results. We first focus on the relationship between loads and LMPs assuming the other five factors remain unchanged; then Section III takes the line flow limits and ramp constraints into account; the influence of generation offer prices is explored in Section VII-C. Future work will investigate the impacts of unit commitment results and the system topology changes on the prices.

Consider the static SCED in the standard MLP form[1]:

$$\text{Primal:} \min\{c^\intercal P_G : AP_G + s = b + WP_D, s \ge 0\} \tag{3a}$$

$$\text{Dual:} \max\{-(b + WP_D)^\intercal y : A^\intercal y = -c, y \ge 0\} \tag{3b}$$

where

$$A = \begin{bmatrix} \mathbf{1}_{n_b}^\intercal \\ -\mathbf{1}_{n_b}^\intercal \\ H \\ -H \\ \mathbf{I}_{n_b} \\ -\mathbf{I}_{n_b} \end{bmatrix}, b = \begin{bmatrix} 0 \\ 0 \\ F^+ \\ -F^+ \\ P_G^+ \\ -P_G^- \end{bmatrix}, W = \begin{bmatrix} \mathbf{1}_{n_b}^\intercal \\ -\mathbf{1}_{n_b}^\intercal \\ H \\ -H \\ \mathbf{0}_{n_b \times n_b} \\ \mathbf{0}_{n_b \times n_b} \end{bmatrix} \tag{4}$$

The load vector P_D is the vector of parameters θ, and the load space $\mathbb{D}$ is the parameter space Θ. Since not every P_D in the load space leads to a feasible SCED problem, $\mathcal{D} \in \mathbb{D}$ denotes the set of all feasible vectors of loads. [13] shows that $\mathcal{D}$ is a convex polyhedron in $\mathbb{D}$.

Definition 1 (Optimal Partition/System Pattern). For a load vector $P_D \in \mathcal{D}$, we could find a finite optimal solution P_G^* and s^*. Let $\mathcal{J} = \{1, 2, \dots, n_c\}$ denote the index set of constraints where $n_c = 2 + 2n_l + 2n_g$ for Eqn. (3). The *optimal partition*

---

[1]In other references (e.g. [12], [13]), the primal form of the MLP problem is different. For the consideration of convenience of analyzing SCED problem, we follow the formulations in [11]. Those two forms are interchangeable.

$\pi = (\mathcal{B}, \mathcal{N})$ of the set $\mathcal{J}$ is defined as follows:

$$\mathcal{B}(P_D) := \{i : s_i^* = 0 \quad \text{for} \quad P_D \in \mathcal{D}\} \tag{5a}$$

$$\mathcal{N}(P_D) := \{j : s_j^* > 0 \quad \text{for} \quad P_D \in \mathcal{D}\} \tag{5b}$$

Obviously, $\mathcal{B} \cap \mathcal{N} = \emptyset$ and $\mathcal{B} \cup \mathcal{N} = \mathcal{J}$. The optimal partition $\pi = (\mathcal{B}, \mathcal{N})$ divides the index set into two parts: binding constraints $\mathcal{B}$ and non-binding constraints $\mathcal{N}$. In SCED, the *optimal partition* represents the status of the system (e.g., congested lines, marginal generators), and is called *system pattern*.

Definition 2 (Critical Region/SPR). The concept *critical region* refers to the set of vectors of parameters which lead to the same optimal partition (system pattern) $\pi = (\mathcal{B}_\pi, \mathcal{N}_\pi)$:

$$\mathcal{S}_\pi := \{P_D \in \mathcal{D} : \mathcal{B}(P_D) = \mathcal{B}_\pi\} \tag{6}$$

For the consideration of consistency, the *critical region* is called SPR in this paper.

According to the definitions, each SPR is one-to-one mapped to a system pattern, the SPRs are therefore disjoint and the union of all the SPRs is the feasible set of vectors of loads: $\cup_i \mathcal{S}_{\pi_i} = \mathcal{D}$. All the SPRs together represent a specific partition of the load space. The features of SPRs, which directly inherit from critical regions in MLP theory, are summarized as follows:

Theorem 1. The load space could be decomposed into many SPRs. Each SPR is a convex polytope. The relative interiors of SPRs are disjoint convex sets and each corresponds to a unique system pattern [6]. There exists a separating hyperplane between any two SPRs [8].

Lemma 1 (Complementary Slackness). According to complementary slackness:

$$A_\mathcal{B} P_G^* = (b + W P_D)_\mathcal{B} \tag{7a}$$

$$A_\mathcal{N} P_G^* < (b + W P_D)_\mathcal{N} \tag{7b}$$

$$A_\mathcal{B}^\mathsf{T} y_\mathcal{B} = -c, y_\mathcal{B} > 0 \tag{7c}$$

$$y_\mathcal{N} = 0 \tag{7d}$$

where the $(\cdot)_\mathcal{B}$ is the sub-matrix or the sub-vector whose row indices are in set $\mathcal{B}$, same meaning applies for $(\cdot)_\mathcal{N}$.

Remark 1: The supply-demand balance equality constraint in is rewritten as two inequalities in Eqn. (3). These two inequalities will always be binding and appear in the binding constraint set $\mathcal{B}$ at the same time. One of them is redundant and therefore eliminated from the set $\mathcal{B}$. *In the remaining part of the paper, set $\mathcal{B}$ denotes the set after elimination.* If the problem is not degenerate, the cardinality of binding constraint set $\mathcal{B}$ is equal to the number of decision variables (i.e., number of generators n_g),[2] the matrix $A_\mathcal{B}$ is invertible.

Remark 2: SCED problems with different generation costs will have different SPRs. For a system pattern $\pi = (\mathcal{B}, \mathcal{N})$, its SPR would remain the same as long as the generation cost vector c satisfies Eqn. (7c).

Lemma 2: Within each SPR, the vector of LMPs is unique [7], [8].

The proof of this lemma follows Eqn. (7c) (dual form of system pattern definition). Since the system pattern π is unique
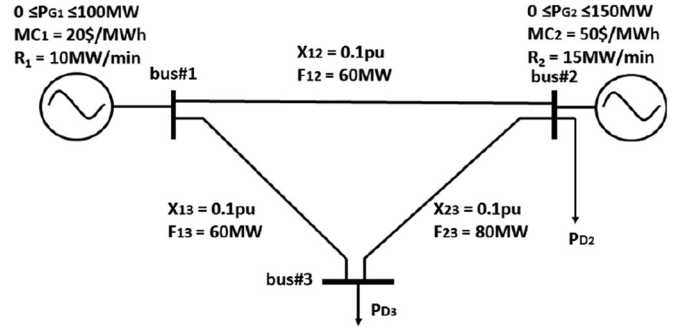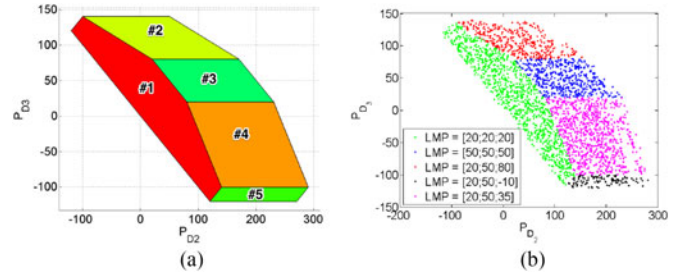

Fig. 1. 3-bus system.

Fig. 2. SPRs of the 3-bus system (static SCED). (a) Theoretical Results Using MPT 3.0 and (b) Monte-Carlo Simulation.

within an SPR S_π, therefore the solution y^* is unique for any $P_D \in \mathcal{S}_\pi$ And the vector of LMPs can be calculated using Eqn. (2). This lemma also illustrates that the LMP vectors are discrete by nature in the case of linear costs.

Theorem 2: If the SCED problem is not degenerate, then different SPRs have different LMP vectors.

The proof of Theorem 2 turns out to be non-trivial, and is described as follows. If two SPRs have the same LMP: $\lambda^{(i)} = \lambda^{(j)}$, their energy components are the same because of the entry-wise equality, then Eqn. (2) suggests that the congestion components should also be the same: $H^\mathsf{T}(\mu^{(i)} - \mu^{(j)}) = 0$. Given the fact that the null space of H^T is always non-empty,[3] a critical question arises: "is it possible that $\mu^{(i)} - \mu^{(j)}$ belongs to the null space of H^T?" Or equivalently, "is it possible that different congestion patterns have the same LMP vector?" We show that the answer is "no". A complete proof of the theorem is provided in [14].

D. An Illustrative Example

The 3-bus system in Fig. 1 serves as an illustrative example in this paper. It was first analyzed using the Multi-Parametric Toolbox 3.0 (MPT 3.0) [15], results are shown in Fig. 2(a). A Monte-Carlo simulation is conducted, with load vectors colored according to their LMPs. The theoretical results are verified by the Monte-Carlo simulation results. Notice that P_{D_2} and P_{D_3} could be negative. This is for the consideration of renewable resources in the system, which are typically considered as negative loads.

---

[2]This is consistent with the statement that the number of marginal generators equals to the number of congested lines plus one.

[3]$\dim(N(H^\mathsf{T})) = n_l - n_b + 1 \geq 0$. The equality holds if and only if the topology of the system is a tree, where $n_l = n_b - 1$.
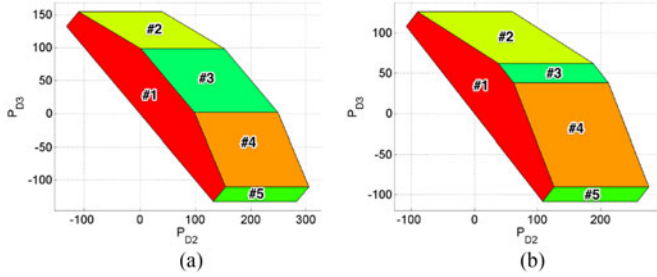
Fig. 3. SPRs of the 3-bus system (static SCED with DLRs). (a) Line Limits: (66, 66, 88) and (b) Line Limits: (54, 54, 72).

III. SPRs WITH VARYING PARAMETERS

Section II-C shows construction properties of the load space with fixed parameters of the system (e.g., transmission constraints). However, these parameters might be time-varying due to reasons like dynamic ratings or active ramping constraints. This subsection reveals more features of SPRs with respect to varying factors in the system.

Lemma 3 (Analytical Form of SPRs). Let $\mathbf{I}_\mathcal{B} \cdot (b + WP_D)$ represent the sub-vector $(b + WP_D)_\mathcal{B}$, where $\mathbf{I}_\mathcal{B}$ is the sub-matrix of the identity matrix whose row indices are in set $\mathcal{B}$. Then the analytical form of the SPRs could be solved from Eqn. (7a) and Eqn. (7b) as follows:

$$(\mathbf{I}_\mathcal{N} A \cdot (\mathbf{I}_\mathcal{B} A)^{-1} \mathbf{I}_\mathcal{B} - \mathbf{I}_\mathcal{N})(b + WP_D) < 0 \qquad (8)$$

Remark 3: Eqn. (8) could be written as:

$$(\mathbf{I}_\mathcal{N} A (\mathbf{I}_\mathcal{B} A)^{-1} \mathbf{I}_\mathcal{B} - \mathbf{I}_\mathcal{N}) \cdot WP_D < (\mathbf{I}_\mathcal{N} - \mathbf{I}_\mathcal{N} A (\mathbf{I}_\mathcal{B} A)^{-1} \mathbf{I}_\mathcal{B})b \qquad (9)$$

This indicates that the shape of the SPR $\mathcal{S}_\pi$ only depends on two factors: (1) the corresponding system pattern $\pi = (\mathcal{B}, \mathcal{N})$; (2) matrices A and W, namely the shift factor matrix H according to Eqn. (4). Small changes of vector b only parallel-shift the SPRs' boundaries.

A. Dynamic Line Rating (DLR)

DLR, contrary to the *static line rating* (SLR), refers to the technology that optimizes the transmission capacity based on the real-time conditions such as ambient temperature and wind speed [16]. It is considered to be more adaptive in maximizing the line potential while keeping the secure grid operation.

From dispatch point of view, DLR can be represented by the changes of transmission limits F^+ in Eqn. (1c). It changes the vector b in Eqn. (4) and thus translate the boundaries of SPRs.

The 3-bus system in Fig. 1 with different transmission limits is analyzed via MPT 3.0. Compared with the standard transmission limits $[60; 60; 80]$, when we increase the limits by 10% (see Fig. 3(a)), SPR #3 expands but SPR #1, #2 and #4 shrink; when we decrease the limits by 10% (see Fig. 3(b)), SPR #3 shrinks but SPR #1, #2 and #4 expand. This verifies the claim that DLRs only shift the boundaries without altering the shapes of SPRs. The implication of having DLR is that SPRs in Fig. 4 are overlapping instead of completely separable in Fig. 2(b). Details of the Monte-Carlo simulation are provided in Section V-C1.
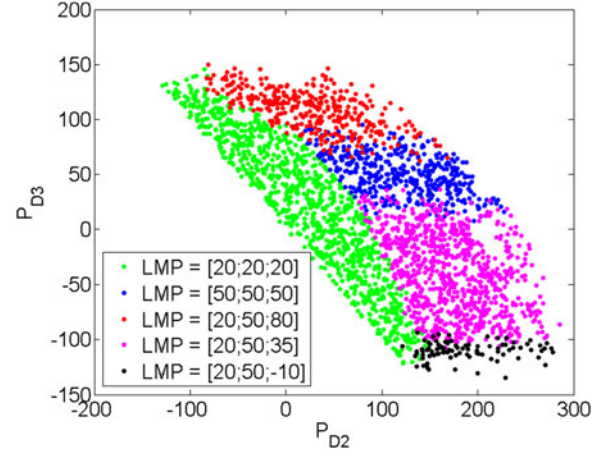

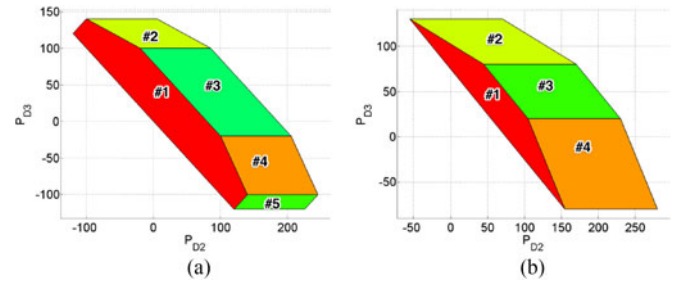Fig. 4. Monte-Carlo simulation (static SCED with DLRs).

Fig. 5. SPRs of the 3-bus system (SCED with ramp constraints). (a) Previous Generation: (30, 30) and (b) Previous Generation: (100, 100).

B. Ramping Constraints

The analysis of SPRs can also be generalized to the dispatch models that include inter-temporal constraints such as ramping:

$$P_G^{k-1} - R^- \Delta t \le P_G^k \le P_G^{k-1} + R^+ \Delta t \qquad (10)$$

In Eqn. (10), R^+ and R^- represent the ramp up and down constraints of generators.

Adding ramp constraints to the static SCED problem is equivalent with replacing the generation capacity constraints Eqn. (1d) with:

$$\max\{P_G^-, P_G^{k-1} - R^- \Delta t\} \le P_G^k \le \min\{P_G^+, P_G^{k-1} + R^+ \Delta t\} \qquad (11)$$

When the ramp capacity is not binding, i.e., $P_G^- > P_G^{k-1} - R^- \Delta t$ and $P_G^+ < P_G^{k-1} + R^+ \Delta t$, the SCED problem is the same as the case where no ramp constraints are considered. The SPRs would be exactly the same as in Fig. 2(a) and (b). However, active ramp constraints change the actual generation constraints, and therefore change the parameter b in Eqn. (4). This leads to parallel shift of the boundaries of SPRs. The impacts of ramping constraints on SPRs is similar with the case of DLRs.

The 3-bus system, again, is analyzed via both MPT 3.0 and Monte-Carlo simulation. Fig. 5(a) and (b) demonstrate the cases where ramp constraints are active. SPRs look similar with parallel changes on the boundaries. When analyzing the load and LMP data, we will again see the overlapping SPRs (see Fig. 6).
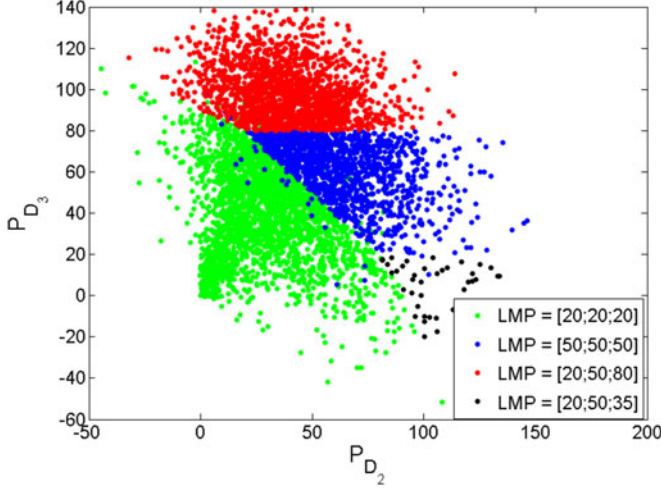
Fig. 6. Monte-Carlo simulation (SCED with ramp constraints).

Fig. 7. SPR identification problem with SVM (separable case).

Fig. 8. SPR identification problem with SVM (non-separable case).

IV. A DATA-DRIVEN APPROACH TO IDENTIFYING SPRs

The SPRs depict the fundamental coupling between loads and LMP vectors. Massive historical data could help market participants estimate SPRs, understand the load-LMP coupling and then forecast LMPs. This section proposes a data-driven method to identify SPRs, which is a significant improvement of the basic method in [8] by considering varying system parameters and the probabilistic nature of system parameters.

A. The SPR Identification Problem

1) SPR Identification as a Classification Problem: A *classifier* is an algorithm to give a *label* y to each *feature* vector x. The feature vectors sharing the same labels belong to the same *class*. The objective of the classification problem is to find the best classifier which could classify each feature vector accurately. For the parametric classifiers, there is always a training set, i.e., a group of feature vectors whose labels are known. There are two steps in a classification problem: training and classifying. *Training* usually means solving an optimization problem over the training set to find the best parameters of the classifier. And *classifying* is to classify a new feature vector with the trained classifier.

According to Section II-C, the load vectors in an SPR share many common features (e.g., vectors of LMPs). Theorem 2 proved that the LMP vectors are distinct for different SPRs. Therefore, one SPR can be regarded as a *class* and the LMP vector is the *label* of each class. Theorem 1 proves the existence of the separating hyperplanes. Since each separating hyperplane labels two SPRs with different sides, it turned out that the separating hyperplanes are classifiers and the key of identifying SPRs is to find optimal hyperplanes, which is exactly the objective of Support Vector Machine (SVM).

2) SPR Identification with SVM: Suppose there is a set of labeled load vectors for training and those load vectors belong to only *two* distinct SPRs (labels $y^{(i)} \in \{1, -1\}$). Then the SPR identification problem with a *binary* SVM classifier (separable
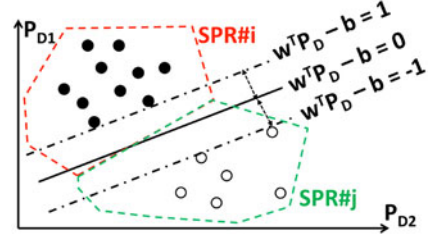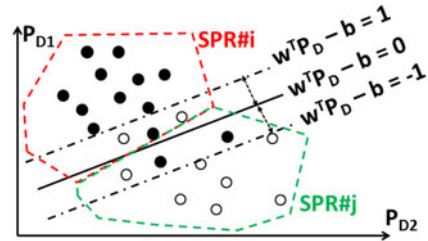
case) is stated below:

$$\min_{w,b} \quad \frac{1}{2} w^\mathsf{T} w \tag{12a}$$

$$\text{s.t} \quad y^{(i)}(w^\mathsf{T} P_D^{(i)} - b) \geq 1, y^{(i)} \in \{-1, 1\} \tag{12b}$$

The word "binary" here specifies only two classes (i.e., SPRs) are being considered. Eqn. (12b) is feasible only when the two SPRs are not overlapping and there exists at least one hyperplane thoroughly separating them (separable case, Fig. 7). For any load vector P_D in the load space, $w^\mathsf{T} P_D - b = 0$ represents the separating hyperplane where w is the norm vector to the hyperplane. Two lines satisfying $w^\mathsf{T} P_D - b = \pm 1$ separate all the training data and formulate an area with no points inside. This empty area is called *margin*. The width of the margin is $2/||w||$, which is the distance between those two lines. The optimal solution refers to the separating hyperplane which maximizes the width of the margin $2/||w||$, therefore the objective of the binary SVM classifier is to minimize the norm of vector w.

Due to the existence of multiple SPRs, multi-class classifiers are needed. Since Theorem 1 guarantees the existence of separating hyperplanes between every pair of SPRs, the "one-vs-one" multi-class SVM classifier is incorporated in the data-driven approach to identifying SPRs. Detailed procedures are summarized in Section IV-B.

B. A Data-Driven Approach

1) SPR Identification with Varying System Parameters: When the system parameters are varying (e.g., DLRs), two SPRs may overlap with each other (Fig. 8). The SPR identification problem is no longer a separable case as in Section IV-A2. The SVM classifier needs to incorporate *soft margins* to allow some tolerance of classification error. The slack variable s_i is added to Eqn. (12a) and penalties of violation $C \sum_i s_i$ are
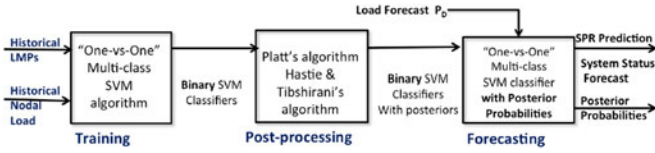
Fig. 9.   The data-driven approach.

added in the objective function. Large C indicates low extent of tolerance.

$$\min_{w,b,s} \quad \frac{1}{2} w^\mathsf{T} w + C \sum_i s^{(i)} \tag{13a}$$

$$\text{s.t} \quad y^{(i)}(w^\mathsf{T} P_D^{(i)} - b) \geq 1 - s^{(i)} \tag{13b}$$

$$s^{(i)} \geq 0, y^{(i)} \in \{-1, 1\}$$

2) Fitting Posterior Probabilities: The posterior probability is the probability that the hypothesis is true given relevant data or observations. In the classification problem, the posterior probability can be stated as: $\mathbb{P}(\text{class}|\text{input})$.

Estimating the posterior probability is very helpful in practical problems [17]. When identifying SPRs, knowing the posterior probability $\mathbb{P}(y = i|P_D \text{ and } y \in \{1, 2, \ldots, n\})$ is not only about knowing the classification result $y = i$ (P_D belongs to SPR#i), but also understanding the confidence or possible risk. The market participants could accordingly adjust their bidding strategy and reduce possible loss.

Although the posterior probabilities are desired, the standard SVM algorithm provides an uncalibrated value which is not a probability as output [17]. Modifications are needed to calculate the *binary* posterior probabilities $\mathbb{P}(y = i|P_D \text{ and } y \in \{i, j\})$. Common practice is to add a link function to the binary SVM classifier and train the data to fit the link function. Some typical link functions include sigmoid functions [17] and Gaussian approximations [18]. In this paper, the sigmoid link function is selected due to its general better performance than other choices [17].

In general, there are more than two SPRs. What we really want to know is the *multi-class* posterior probabilities $\mathbb{P}(y = i|P_D \text{ and } y \in \{1, 2, \ldots, n\})$. For short, we will use $\mathbb{P}(y = i|P_D)$ to represent multi-class posterior probabilities. [18] proposed a well-accepted algorithm to calculate multi-class posterior probabilities from pairwise binary posterior probabilities. This algorithm is incorporated in our approach and briefly summarized in our technical report [14].

3) The Data-Driven Approach: There are three steps in the proposed data-driven approach (see Fig. 9):

a) Training: Suppose there are n different SPRs in the training data set. Each time two SPRs are selected, trained and we get a *binary* SVM classifier. This pairwise training procedure is repeated $C_n^2 = n(n-1)/2$ times and we collect $n(n-1)/2$ binary classifiers, namely the $n(n-1)/2$ separating hyperplanes between any two out of n SPRs.

b) Classifying/Predicting: Given load forecast P_D, we could use the "*max-vote-wins*" algorithm to get the classification results: each binary classifier provides a classification result (vote) for the load forecast P_D, the SPR which collects

the most votes will be the final classification result. The load forecast P_D is therefore pinpointed to an SPR. The LMP forecast: $\hat{\lambda}(P_D) = \lambda^{(i^*)}$ where i^* is the index of the SPR winning the most votes. This step is independent of the data post-processing procedure.

c) Data Post-Processing: Calculate posterior probabilities $\mathbb{P}(y = i|P_D)$ for $i = 1, 2, \ldots, n$ by applying Platt's algorithm and then Hastie and Tibshirani's algorithm.[4] It is worth noting that the proposed approach is generalizable to many other scenarios with overlapping SPRs in the data, possible extensions are discussed in Section VII-A.

V. CASE STUDY

In this section, we illustrate the proposed data-driven approaches on two systems.

A. Performance Metrics

We first introduce the performance metrics.

1) 5-fold Cross Validation: To evaluate the performance of the model to an independent data set and avoid overfitting, the *k-fold cross validation* technique is being used. In *k*-fold cross-validation, the overall data set is randomly and evenly partitioned into k subsets. Every time a subset is chosen as validation data set, and the remaining $k - 1$ subsets are used for training. This cross-validation process is repeated k times (k folds), and each subset serves as the validation data set once. The 5-fold cross validation is being used in this paper.

2) Classification Accuracy: Classification accuracy is the most common criteria to evaluate the performance of classifiers. The classification accuracy α is the ratio of the correctly classified points in the validation data set. When incorporating 5-fold cross validation, the classification accuracy of each fold $(\alpha_1, \alpha_2, \ldots, \alpha_5)$ is calculated first, then the overall performance of the method is evaluated by the average classification accuracy $\bar{\alpha} = (\sum_{i=1}^{5} \alpha_i)/5$.

3) LMP Forecast Accuracy: The proposed approach forecasts the LMP at every bus. The performance of LMP forecast at bus i is evaluated by the nodal LMP forecast accuracy β_i, which is the average forecast accuracy of all the validation data points ($j = 1, 2, \ldots, n_v$)

$$\beta_i = \frac{1}{n_v} \sum_{j=1}^{n_v} \frac{|\hat{\lambda}_i[j] - \lambda_i[j]|}{\lambda_i[j]} \tag{14}$$

The overall LMP forecast accuracy β evaluates the performance of LMP forecast for the whole system. It is the average of all the nodal LMP forecast accuracy β_i ($i = 1, 2, \ldots, n_b$):

$$\beta = \frac{1}{n_b} \sum_{i=1}^{n_b} \beta_i \tag{15}$$

B. Static SCED With SLRs

This section explores the simplest case: static SCED with SLRs. Since [8] discusses the 3-bus system as well as the IEEE

---

[4]Details of these two algorithms are summarized in [14].

TABLE I
RESULTS OF THE 118-BUS SYSTEM (STATIC SCED WITH SLR)

Fold	Classification	LMP Forecast
1	64.24%	96.82%
2	67.36%	96.71%
3	64.93%	96.95%
4	71.18%	97.34%
5	65.63%	96.84%
Avg	66.67%	96.93%

TABLE II
RESULTS OF 3-BUS SYSTEM (FIVE-FOLD VALIDATION)

Fold	Classification	LMP Forecast
1	93.967%	96.218%
2	93.236%	96.054%
3	94.150%	95.767%
4	95.612%	96.700%
5	94.150%	96.405%
Avg	94.23%	96.23%

24-bus system, we only examine the data-driven approach on an 118-bus system. The same dataset generated in this section is used in Section VI-A as well.

1) System Configuration: Most of the system settings follow the *IEEE 118-bus, 54-unit, 24-hour system* in [19] but with the following changes: (1) the lower bounds of generations are set to zero, but the upper bounds of generators remain the same as in [19]; (2) generation costs are linear. Details of the parameters are summarized in [20].

2) Load: [19] also provides an hourly system load profile and a bus load distribution profile. With linear interpolation, the hourly system load profile is modified to be 5-min based. To account for the variability of loads, we assume the load at each bus follows normal distribution $\mathcal{N}(\mu, \sigma)$. The expectation μ of each nodal load is calculated from the system load profile and bus load distribution profile, the standard deviation σ is set to be 10% of the expectation. 1440 (5 days, 5-min based) load vectors are generated, then Matpower [21] solves these 1440 SCED problems and records 1440 LMP vectors. These 1440 load vectors and LMP vectors are the training and validation data.

3) Simulation Results: Results are summarized in Table I. The classification accuracy is around 67% but the LMP forecast is satisfying. When the classification result of a load vector is correct, the LMP forecast is correct for every bus, i.e., $\beta = 100\%$. It is worth noting that even if the classification fails, the overall LMP forecast still has accuracy about 90%. This is because the classification errors happen between one SPR and its neighbors. LMPs of adjacent SPRs are similar due to the fact that only one active constraint is different.[5] Therefore, the LMP forecast result is much more accurate than classification.

C. Static SCED with DLRs

1) 3-bus System: We start with an illustrative 3-bus system example. This succinct example provides key insights and visualization of the proposed method.

a) Data: The parameters of the 3-bus system are presented in Fig. 1. The data set is generated using Matpower with the following assumptions: (1) the load vector is evenly distributed in the load space; and (2) the transmission limits F is time-varying: for simplicity, we utilize the following model to calculate the real-time transmission limits F:

$$F = (1 + \xi)F_0 \tag{16}$$

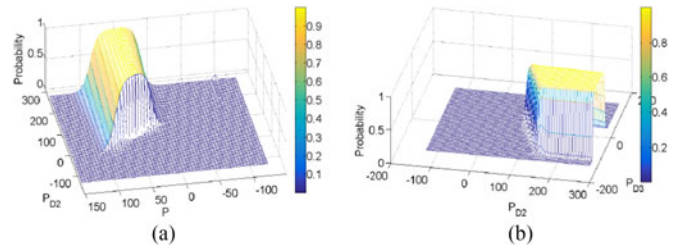[5]Lemma *System Patterns of Adjacent SPRs* in [14].

Fig. 10. Posterior probabilities of two SPRs. (a) SPR#3: LMP = (50, 50, 50) and (b) SPR#4: LMP = (20, 50, 35).
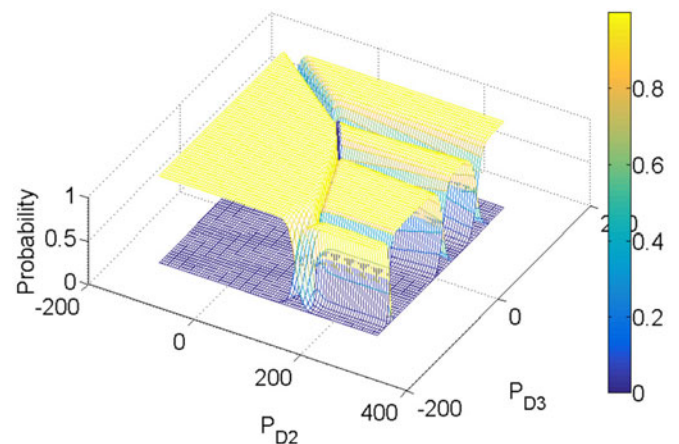

Fig. 11. Posterior probability surfaces.

F_0 is the "standard" transmission limits and $F_0 = [60; 60; 80]$. It is the same as the case of SLRs. $\xi \sim N(0, 0.1)$ represents the major factor (e.g., ambient temperature or wind speed) that impacts the transmission capacities. All the data generated is visualized in Fig. 4.

b) Simulation results: Table II summarizes the classification and LMP forecast accuracies. The accuracies are around 95% because of the overlapping SPRs.

c) Posterior probabilities: The posterior probabilities are visualized. The posterior probabilities of an SPR compose a surface (see Fig. 10(b)). When putting all the 5 surfaces of 5 SPRs together (shown in Fig. 11), the five surfaces intersect with each other and formulate some "mountains" and "valleys". The "mountains" correspond to the inner parts of SPRs, where the overlapping of SPRs is almost impossible to happen. And the "valleys" always locate at the boundaries among SPRs.

TABLE III
AVERAGE COMPUTATION TIME (IN SECONDS)

Steps	Computation Time (s)
(a) Training	58.73
(b) Predicting (288 points)	26.8504
(c) Data post-processing	701.22

TABLE IV
RESULTS OF THE 118-BUS SYSTEM (DLR)

Fold	Classification	LMP Forecast
1	61.11%	95.11%
2	59.38%	94.53%
3	60.76%	95.24%
4	51.39%	93.34%
5	55.90%	94.22%
Avg	57.71%	94.49%

TABLE V
RESULTS ON 118-BUS SYSTEM WITH DIFFERENT RAMP SETTINGS

R/R_0	0.5	1	2
LMP Forecast	44.57%	85.10%	96.33%

2) 118-bus System: A more comprehensive case study is conducted on the 118-bus system to evaluate the performance and computational burden of the data-driven approach on a complex system with realistic settings.

a) System configuration: The only difference from the system configuration in Section V-B is about transmission limits. To consider DLR, we use the same model as Eqn. (16). F_0 is the same as the transmission limits in [19] and $\xi \sim N(0, 0.1)$.

b) Performance: The algorithm is implemented using the Statistics and Machine Learning Toolbox of Matlab. Simulation results are presented in Table IV, and Table III summarizes the computation time of each step in the data-driven approach on a PC with Intel i7-2600 8-core CPU@3.40GHz and 16GB RAM memory. There are 181 SPRs found in 1152 points for training, $C_{181}^2 = 16290$ SVM classifiers are trained in 58.72 seconds. On average, one SVM classifier is trained within 0.004 seconds. This is because most of the SPRs are completely separable, these cases will be solved in an extremely short time. Those adjacent SPRs are overlapping and are the major source of the computational burden.

D. Case Studies With Ramp Constraints

1) Settings: The parameters of the 118-bus system are the same as in Section V-B. And the ramp capacities of generators follows the simplified assumption below: each generator could ramp up (down) to its generation limits in 15 min. For example, a generator with $G^+ = 200$ MW and $G^- = 125$ MW, its ramp capacity is: $R^+ = R^- = (200 - 125)/15 = 5$ MW/min. This setting is called R_0 in Table V. Due to the temporal cou-

pling of SCED with ramp constraints, a daily load curve is necessary. The settings of loads are the same as in Section V-B. 1440 SCED problems are solved consecutively with Matpower, and 1440 load vectors and LMP vectors are recorded.

2) Simulation Results: The classification and LMP forecast accuracies are summarized in Table V. With the ramp settings above, the classification and LMP forecast are satisfying. However, different ramp settings would change the results dramatically. As shown in Table V: when generators ramp up/down 2 times faster ($R/R_0 = 2$), the ramp constraints would rarely be active, then it is the same as static SCED; when generators ramp up/down 2 times slower ($R/R_0 = 0.5$), the actual generation upper/lower bounds are determined by previous dispatch results and ramp constraints. Generation limits become time-varying and the SPRs are overlapping. This explains the unsatisfying results when the system is lack of ramp capacities. Furthermore, varying SPRs could also explain the price spikes during ramping up hours in the morning and ramping down hours in the early evening.

VI. THE IMPACT OF NODAL LOAD INFORMATION

We would like to point out that one possible contribution of this paper is to consider the LMP changes due to nodal load variations. This section dedicates to a detailed discussion about the impact of nodal load information on the understanding of LMP changes. We first demonstrate the benefits of having nodal load information in Section VI-A; then Section VI-B illustrates the effects of incomplete load information and the attempts to solve the issue.

To concentrate on the effects that incomplete load information brings, we make the following assumptions: (1) transmission limits are constant, no DLRs are being considered; (2) ramp constraints are not taken into account.

A. On Nodal Load Levels

Previous literature such as [2] studied the impact of system load levels on the LMPs. An important concept "CLL" is defined as the system load level where the step changes of LMPs happen. Many LMP forecast methods were proposed based on identifying CLLs. But the definition of CLL assumes that the nodal load levels of all the buses change *proportionally*. This assumption constrains the load vectors in the load space to be on a straight line, and the CLLs are indeed the intersection points of the straight line with the boundaries of SPRs.

We would like to point out that one possible contribution of this paper is to consider the LMP changes due to *nodal-level* load variations. Contrary to CLL-based methods, which solve a one-dimensional problem, the proposed SVM-based method could explore all the dimensions of the load space and is indeed a generalization of the CLL-based method.

Consider the SPR identification problem with only one feature vector: the total demand of the system. Fig. 12 illustrates the problem formulation. Since only the total demand $P_D = P_{D_1} + P_{D_2}$ is available, the load vectors in the original SPRs are projected to the axis of total demand. Because this is a one-dimension problem, the SVM classifier degenerates to the
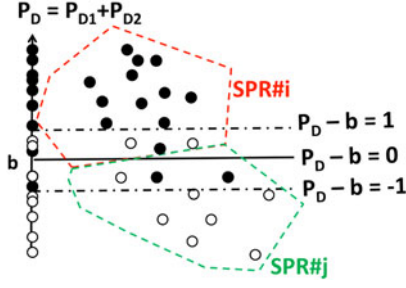
Fig. 12. Identifying CLLs.

TABLE VI
COMPARISON OF CLL AND SVM (118-BUS SYSTEM)

LMP Forecast	CLL	SVM
Overall	94.82%	95.95%
Price > 45 $/MWh	88.86%	96.32%
Worst Forecast (Bus No.)	73.92% (23)	88.17% (23)

TABLE VII
RESULTS 3-BUS SYSTEM

LSE	LMP@Bus 1	LMP@Bus 2	LMP@Bus 3	Overall
1	86.08%	97.45%	88.53%	90.69%
2	70.69%	96.13%	89.31%	85.38%
3	87.91%	98.65%	93.53%	93.53%
CLL	69.48%	97.15%	89.24%	85.29%

case that there is only one decision variable b, the direction of the separating hyperplane w is represented by the positivity of b. The objective becomes finding the optimal value b which has the least overlapping points of different LMPs

$$\min_{b,s} \quad \sum_i s^{(i)} \tag{17a}$$

$$\text{s.t} \quad y^{(i)}(P_D^{(i)} - b) \geq 1 - s^{(i)} \tag{17b}$$

$$s^{(i)} \geq 0, y^{(i)} \in \{-1, 1\}$$

We compare this CLL-based method and SVM-based method on the 3-bus system and 118-bus system. Results are demonstrated in Tables VI, VII, and Fig. 13. The performance of both methods are close for the nodal LMP forecast of many buses, but the CLL-based method failed to provide correct forecast of some specific buses (e.g., bus 23 in Fig. 13), while the SVM-based method provides much better results. The SVM-based method is also better on forecasting high prices.

B. Incomplete Load Information

In practice, LSEs or other market participants may not have the complete information about load levels at all buses. We investigate the performance of the algorithm by assuming LSEs have access to only: (1) the total system-level load; and (2) the nodal load levels in its own area.
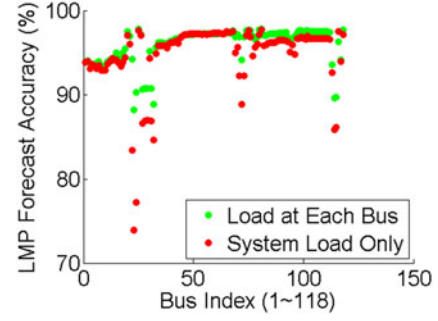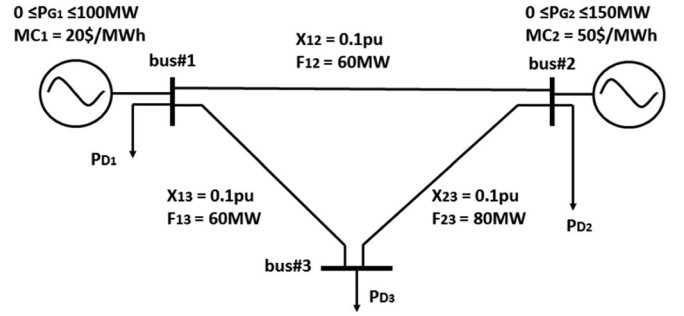

Fig. 13. Nodal LMP forecast accuracy.
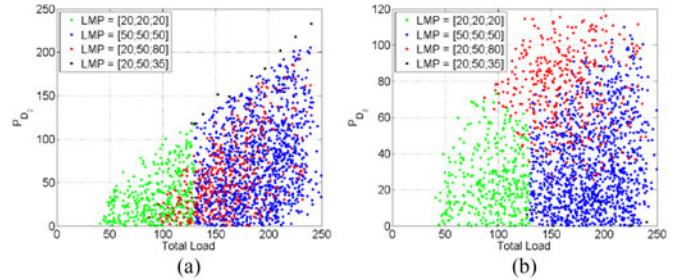

Fig. 14. 3-bus system with 3 loads.

Fig. 15. LSEs. (a) LSE 2 and (b) LSE 3.

To better illustrate the problem formulation. We add a load P_{D_1} at bus 1 to the 3-bus system in Fig. 1.[6] Modified system is shown in Fig. 14. Assume there are three LSEs in the system. LSE #i at bus i has access to the following information: (1) load at bus i: P_{D_i}; and (2) system-level load: $\sum_{i=1}^{3} P_{D_i}$.

With incomplete load information, the SPR identification problem becomes more difficult. For example, LSE 2 observes two SPRs which almost completely overlap with each other (blue and red in Fig. 15(a)). Since the one-to-one mapping of the LMP vectors and SPRs is not effected by the incomplete load information, this is still a classification problem. The data-driven approach can still be applied but the feature vectors are the system load and a subset of nodal load levels, instead of load levels at every bus in Section IV.

---

[6]If there are still two loads in the system, knowing system-level load $P_{D_2} + P_{D_3}$ and P_{D_2} is equivalent with knowing P_{D_2} and P_{D_3}.
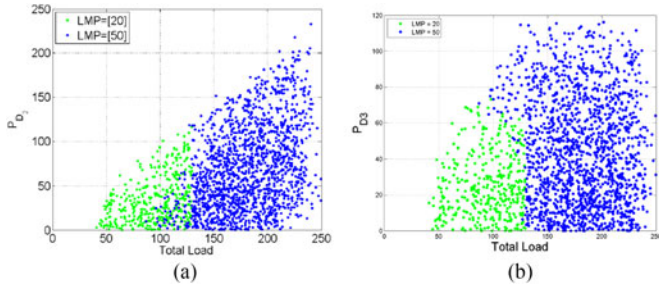
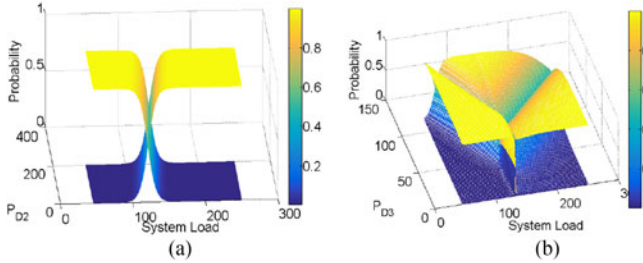Fig. 16.   LMP at bus 2. (a) P_{D_2} and system load and (b) P_{D_3} and system load

Fig. 17.   Posterior probability surfaces. (a) LSE 2 and (b) LSE 3

Simulation results are summarize in Table VII. The results indicate that classification accuracy goes down to around 50% while the LMP forecast accuracy is still satisfactory. This could be explained by the following observations: (1) Fig. 15(a) and (b) is obtained by projecting the 3D SPRs to a lower dimension space. Since the projection is a linear transformation, although the SPRs are overlapping, their boundaries remain linear; (2) the LSEs may care more about their own LMPs. For example, Fig. 15(a) could be re-colored by the LMPs at bus 2 (see Fig. 16(a)). Since there are only two possibilities of LMPs at bus 2 (20 and 50), there are only two colored regions in Fig. 16(a). Even with relatively low accuracy of the overall classification, the forecast of LMPs at bus 2 is still accurate.

When forecasting a subset of nodal LMPs becomes the major concern, it might be more computationally efficient to formulate the problem in a way as Fig. 16(a) shows. The number of classes decreases significantly and so does the computational burden. But the new colored regions might be the union of SPRs. Though the colored regions in Fig. 16(a) are convex, the union of convex sets are usually non-convex. Because of this, the SVM with linear kernel may not be the best choice. Choosing the best classifier would depend upon the feature of the regions, and will be part of the future work.

Similar with the case of DLRs or ramp constraints, overlapping SPRs implies uncertainties and the posterior probabilities are necessary. In Fig. 17, the posterior probabilities for LSE #2 and #3 are visualized, respectively.

Because of the relatively small resistances of transmission lines, the loss components of LMPs are usually small compared with the other two components. Geometrically speaking, each LMP vector is a point in the LMP space and the LMPs of the same SPR form a cluster. The center of the cluster contains the energy, congestion component and the average loss component, the deviations from the center represent varying loss

TABLE VIII
NUMBER OF SPRs OF SOME BENCHMARK SYSTEMS

System Info	MPT 3.0	Simulation (8640 points)
3 Bus System (Fig. 1)	5	4
IEEE 6 Bus System	20	7
IEEE 9 Bus System	15	7
IEEE 14 Bus System	1470	50
IEEE 24 Bus System	$\sim 10^6$	445
IEEE 118 Bus System	–	971

components due to different line flows. We could run a clustering algorithm (e.g., K-means) on the LMP data in order to find out the centers of those clusters. Then the LMP vectors of the same cluster are regarded as the LMPs of the same SPR. By doing so, the SPR identification problem is modeled as a classification problem. The LMP forecast is the forecast of energy components, congestion components and the average loss components.

VII. Discussions

A. On Posterior Probabilities

When dealing with uncertainties, it is natural to analyze the data in a probabilistic manner. The calculation of posterior probabilities is essential and provides the quantification of possible risks. We only propose the method to calculate posterior probabilities in this paper, but quantification of the posterior probabilities could yield many interesting applications. For example, LSEs could consider demand response mechanisms to partially change the load vector and thus shift from high price SPRs. Market participants could also estimate the price volatilities due to renewables in a system. Further discussions on how to utilize the posterior probabilities for specific applications are our future work.

B. On the Computational Cost

The theoretical analysis reveals that the load space could be partitioned into many SPRs. This overall structure of the load space could help solve the SCED problem and shift part of the online computational burden to offline [22]. The total number of SPRs could help evaluate computational burden to some extent.

With MPT 3.0, the exact number of SPRs of some IEEE benchmark systems are calculated (Table VIII). Though the total number of SPRs is finite,[7] it grows extremely fast with the scale of the system. However, with the Monte-Carlo simulation, we found much less SPRs than the theoretical results. Zhou *et al.* [6] point out that because of the regular patterns of loads, only some subsets of the complete theoretical load space could be achievable thus helpful in practice. Therefore, only a small subset of the SPRs is meaningful to be analyzed. This suggests the great potential of reducing the computational burden. The proposed approach is also parallel computation-friendly, which could be very useful when dealing with large-scale simulations.

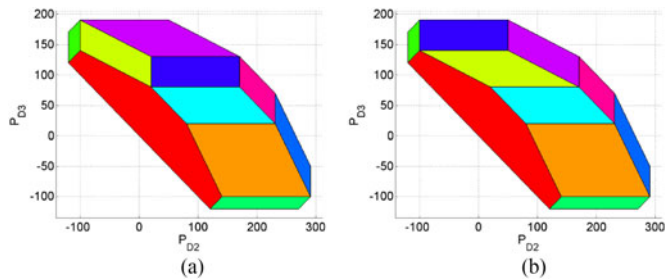[7]A loose upper bound is $2^{n_g-1} \times C_{n_g+n_l}^{n_g-1}$.

Fig. 18. SPRs with different generation offer prices. (a) $c = (20, 50, 65)$ and (b) $c = (20, 50, 100)$.

C. On Generation Offer Prices

The marginal costs of generators are fluctuating due to many factors such as oil prices. This leads to the changes of generation offer prices c in the SCED formulation. Intuitively, the SPRs would change with respect to large offer price variations. Eqn. (7c) in Lemma 1 quantifies the variation of offer prices: for a system pattern $\pi = (\mathcal{B}, \mathcal{N})$, the corresponding SPR $\mathcal{S}_\pi$ would remain the same as long as the generation cost vector c satisfies Eqn. (7c).

An illustrative example is provided below. Suppose a diesel turbine is added at bus 3 in Fig. 1, and the offer price of the diesel turbine is varying due to the fluctuations of oil prices. Fig. 18(a) shows the SPRs when the offer price of the new generator is 65; when the offer price increases from 65 to 100, three SPRs are different while the others remain the same.[8] This shows that the SPRs have some extent of robustness to the varying generation offer prices.

D. LMPs With Loss Components

Since the line losses are not explicitly modeled in the SCED formulation, all the theoretical analysis is conducted on the lossless LMP vectors. The LMP forecast discussed above is the forecast of the energy components and congestion components. In the reality, the proposed method could be applied directly on the markets not considering line losses (e.g., ERCOT), and the electricity markets providing the energy component, congestion component and loss component separately (e.g., MISO). There are many possible methods to forecast the loss components, but that is a separate story.

There are economic dispatch models with line losses explicitly modeled (e.g., [23]), similar analysis using MLP theory could be conducted but it is beyond the scope of this paper.

VIII. CONCLUSION

In this paper, we examine the fundamental coupling between nodal load levels and LMPs in real-time SCED. It is shown that the load space can be partitioned into convex SPRs, which are one-to-one mapped with distinct LMP vectors. Based on

the theoretical results, we propose a data-driven learning algorithm for market participants to identify SPRs. Identifying SPRs is modeled as a classification problem, and the proposed data-driven approach is built upon a "one-vs-one" multi-class SVM classifier. The proposed algorithm is shown to be capable of estimating SPRs solely from historical data without knowing confidential system information such as network topology and bidding curves. The approach is shown to be extensible towards considering DLRs, line losses and partial load information. Simulation results based on the IEEE 118-bus system demonstrates that the proposed algorithm is effective in understanding the past and predicting the future.

This paper is a first step towards developing theoretically rigorous and computationally feasible algorithms to analyzing the market prices as a result of varying loading levels. Future work should investigate: (1) the SPRs with different unit commitment results and system topologies; (2) the impacts of multi-interval temporal constraints on the SPRs. Another important avenue of research is to develop efficient learning algorithm to process a large amount of historical data in near real-time market operations.

REFERENCES

[1] A. Conejo and E. Castillo, "Locational marginal price sensitivities," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 2026–2033, Nov. 2005.
[2] F. Li, "Continuous locational marginal pricing (CLMP)," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1638–1646, Nov. 2007.
[3] F. Li and R. Bo, "Congestion and price prediction under load variation," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 911–922, May 2009.
[4] R. Bo and F. Li, "Probabilistic LMP forecasting considering load uncertainty," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1279–1289, Aug. 2009.
[5] R. Bo and F. Li, "Efficient estimation of critical load levels using variable substitution method," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2472–2482, Nov. 2011.
[6] Q. Zhou, L. Tesfatsion, and C. Liu, "Short-term congestion forecasting in wholesale power markets," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2185–2196, Nov. 2011.
[7] Y. Ji, R. Thomas, and L. Tong, "Probabilistic forecast of real-time LMP via multiparametric programming," in *Proc. 48th Annu. Hawaii Int. Conf. Syst. Sci.*, 2015, pp. 2549–2556.
[8] X. Geng and L. Xie, "A data-driven approach to identifying system pattern regions in market operations," in *Proc. IEEE Power & Energy Soc. General Meeting*, 2015.
[9] D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*. New York, NY, USA: Wiley, 2005.
[10] F. Wu, P. Varaiya, P. Spiller, and S. Oren, "Folk theorems on transmission access: Proofs and counterexamples," *J. Regul. Econ.*, vol. 23, pp. 5–23, 1996.
[11] F. Borrelli, A. Bemporad, and M. Morari, "Geometric algorithm for multiparametric linear programming," *J. Optim. Theory Appl.*, vol. 118, no. 3, pp. 515–540, 2003.
[12] I. Adler and R. D. C. Monteiro, "A geometric view of parametric linear programming," *Algorithmica*, vol. 8, no. 1–6, pp. 161–176, Dec. 1992.
[13] T. Gal and J. Nedoma, "Multiparametric linear programming," *Manage. Sci.*, vol. 18, no. 7, pp. 406–422, 1972.
[14] X. Geng and L. Xie, "Learning the LMP-load coupling from data: A support vector machine based approach," arXiv:1603.07276, 2016. [Online]. Available: http://arxiv.org/abs/1603.07276
[15] M. Herceg, M. Kvasnica, C. Jones, and M. Morari, "Multi-parametric toolbox 3.0," in *Proc. Eur. Control Conf.*, 2013, pp. 502–510.
[16] D. A. Douglass and A.-A. Edris, "Real-time monitoring and dynamic thermal rating of power transmission circuits," *IEEE Trans. Power Del.*, vol. 11, no. 3, pp. 1407–1418, Jul. 1996.
[17] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classifiers*, vol. 10, pp. 61–74, 1999.

---

[8] More specifically, we can calculate the condition from Eqn. (7c): if the offer price of the new generator satisfies $c_3 < 2c_2 - c_1 = 80$, then the SPRs in Fig. 18(a) would remain the same

[18] T. Hastie, R. Tibshirani, and Others, "Classification by pairwise coupling," *Ann. Stat.*, vol. 26, no. 2, pp. 451–471, 1998.

[19] *I. I. of Technology*, "IEEE 118-bus, 54-unit, 24-hour system." [Online]. Available: http://motor.ece.iit.edu/data/JEAS_IEEE118.doc

[20] X. Geng, "Understand LMP-load coupling from a market participant perspective: Theory, examples and an SVM-based data-driven approach," Master's thesis, Texas A&M Univ., College Station, TX, USA, 2015.

[21] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[22] Y. Ji, L. Tong, and R. J. Thomas, "Probabilistic forecast of real-time LMP and network congestion," arXiv:1503.06171, 2015.

[23] E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi, "Marginal loss modeling in LMP calculation," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 880–888, May 2004.

Le Xie (M'10–SM'16) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004. He received the S.M. degree in engineering sciences from Harvard University, Cambridge, MA, USA, in June 2005. He received the Ph.D. degree from the Electric Energy Systems Group from the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA, in 2009. He is currently an Associate Professor in the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA, where he is affiliated with the Electric Power and Power Electronics Group. His research interests include modeling and control of large-scale complex systems, smart grid applications in support of renewable energy integration, and electricity markets.
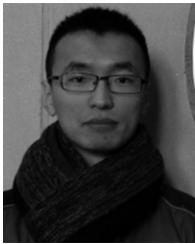
Xinbo Geng (S'14) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013. He received the Master of Science degree from Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA, in 2015. He is currently working toward the Ph.D. degree at Texas A&M University. His research interests include demand response, electricity markets and data analytics in power systems.