

Online Multi-stage Decisions for Robust Power-Grid Operations under High Renewable Uncertainty

Shizhen Zhao, Xiaojun Lin, Dionysios Aliprantis, Hugo N. Villegas
School of ECE,
Purdue University, West Lafayette, IN, USA
Email: {zhao147, linx, dionysios, hvillega}@purdue.edu

Minghua Chen
Department of Information Engineering,
The Chinese University of Hong Kong
Email: minghua@ie.cuhk.edu.hk

Abstract—In this paper, we are interested in online multi-stage decisions to ensure robust power grid operations under high renewable uncertainty. We jointly consider both the reliability assessment commitment (RAC) and the real-time dispatch problems. We first focus on the real-time dispatch problem and define “maximally robust algorithms,” which can provably ensure grid safety whenever there exists any other algorithm that can ensure grid safety under the same level of future uncertainty. We characterize a class of maximally robust algorithms using the concept of “safe dispatch set,” which also provides conditions for verifying grid safety for RAC. However, in general such safe dispatch sets may be difficult to compute. We then develop efficient computational algorithms for characterizing the safe dispatch sets. Specifically, for a simpler single-bus two-generator case, we show that the safe dispatch sets can be exactly characterized by a polynomial number of convex constraints. Then, based on this two-generator characterization, we develop a new solution for the multi-bus multi-generator case using the idea of *virtual demand splitting* (VDS), which can effectively compute a suitable subset of the safe-dispatch set. Our numerical results demonstrate that a VDS-based economic dispatch algorithm outperforms the standard economic dispatch algorithm in terms of robustness, without sacrificing economy.

I. INTRODUCTION

The high variability and uncertainty of renewable energy poses an immense challenge to the existing power grid. Note that in a power grid, the demand and supply must be balanced at all times. Otherwise, the grid will reach an unsafe state. Thus, with high penetration of renewable energy, the grid needs to prepare a sufficient amount of other resources (e.g., traditional fossil-fuel generation) to compensate for the variability and uncertainty of the renewable supply. However, these generation resources and the transmission grid have their own physical constraints, which cannot be violated. Hence, it becomes extremely challenging to determine which set of resources needs to be procured before-hand, and how to dispatch these resources in real time, in order to ensure demand-supply balance at all times subject to the various physical constraints.

In most parts of the US, the responsibility of maintaining grid safety at all times is on the Independent Systems Operator (ISO) [1][2][3]. An ISO typically runs (at least) two markets. In the day-ahead market, a unit-commitment and economic-dispatch schedule is computed for every hour of the next day, based on some forecast of the future demand and renewable supply. Then, in the real-time market during the following day, the ISO must adjust the dispatch decisions every 5 minutes to match the actual demand and supply, which may have deviated from their day-ahead forecasts. In-between the real-

time market and the day-ahead market (or even multiple times during the following day), the ISO runs a RAC¹ (Reliability Assessment Commitment) stage to determine whether additional generators need to be committed so that it has sufficient resources to compensate for future uncertainty in the real-time market. Note that these commitment decisions must be made before-hand because most generators need a substantial amount of lead time to start or stop. In this work, we are most interested in the decisions for both the RAC stage and the real-time dispatch stage, in order to ensure grid safety under high penetration of renewable energy.

Most existing approaches for ensuring reliable grid operations, however, are not suitable for dealing with renewable uncertainty. For example, a majority of the studies, e.g., [4] and [5], have focused on ensuring the so-called $N-k$ reliability upon contingency events, when any k large generators or transmission lines fail. However, renewable uncertainty is present at all times, and thus is intrinsically different from rare contingency events. Recently, two-stage stochastic or robust optimization has been used to study operations under renewable uncertainty [6][7][8][9]. In two-stage optimization, there is a second stage where the future renewable energy supply for all time slots is assumed to be known. In practice, however, the uncertainty of the renewable supply is sequentially revealed as time evolves. Thus, there are naturally many stages where decisions must be made in an online manner. As a result, not only that the real-time dispatch decisions produced under such a two-stage assumption cannot be executed in practice, the decision is also too optimistic for the RAC stage, i.e., it may incorrectly identify a system as safe, even though the system is unsafe [10].

In this work, we aim to account for the multi-stage nature of renewable uncertainty in the RAC and the real-time dispatch decisions. We note that dynamic decision and control problems have been studied in many communication/networking settings. However, it is non-trivial to design solutions that can provide strong safety guarantees in a computationally efficient manner. If one assumes a probabilistic model of the future uncertainty, then this problem can be solved by dynamic programming [11]. However, such an approach suffers from the *curse of dimensionality* when the problem size is large. Lyapunov-based stochastic optimization does not assume any probabilistic model, and has been very successful in many communication/networking problems [12]. However, when faced with tight safety constraints, this approach can only

¹Depending on the ISOs [3], this stage may also be called RUC (Reliability Unit Commitment)

provide a weaker notion of *order-optimality*. In contrast, in this paper we assume that the renewable uncertainty is in a bounded set, and we are interested in designing causal online algorithms that can ensure safe operations of the power grid at all times, as long as the renewable uncertainty is in this uncertainty set. Thus, our approach is at the intersection of robust optimization [13] (which uses bounded uncertainty sets and typically neglects multi-stage decisions) and online algorithms [14] (which consider multi-stage decisions but typically are not designed for a given uncertainty set).

Towards this end, we introduce the concept of “safe dispatch sets” (see Sec. III for the rigorous definition): at the RAC stage, we only need to verify that the safe dispatch sets are non-empty; at the real-time stage, a robust online algorithm only needs to control the dispatch decision to be within the safe dispatch set. We argue that such decisions produce the “most robust” system (see Sec. III for details). Thus, once these safe-dispatch sets are computed, both the RAC problem and the real-time dispatch problem are solved. However, for general settings, computing these safe dispatch sets incurs high computational complexity. Our main contribution is to develop in Sec. IV efficient computational algorithms for characterizing these safe dispatch sets. Specifically, for a simpler single-bus two-generator case, we show that the safe dispatch set can be accurately characterized by a polynomial number of convex constraints. Then, based on this two-generator characterization, we develop a new solution for the multi-bus multi-generator case using the idea of *virtual demand splitting* (VDS), which can effectively compute a suitable subset of the safe-dispatch set.

Our study is closely related to the work in [10] that studies multi-stage robust optimization for power-system operations. In fact, our approximation method in Section IV-B of restricting to a pre-computed *demand splitting factor* may also have some similarity to the restriction to affine real-time dispatch policies in [10]. However, there is one key difference. In robust optimization, the goal is to minimize the worst-case future cost of generation. Thus, only for the worst-case input, one can claim that the affine real-time dispatch policy found in [10] is better. However, the worst case only occurs rarely. For most other future inputs, the economy of such a real-time policy could be poorer than other policies. In contrast, in our solution, we only use the pre-computed splitting factors to determine the safe-dispatch set. Once the safe-dispatch set is found, any decision (in particular, the most economic dispatch decision) in the set can be used (see Section IV-B4 for more details). As we illustrate in the numerical results in Sec. V, our proposed online solution allows the grid operator to balance both reliability and economy, which we believe is highly-appealing in practice.

II. SYSTEM MODEL

We study a power grid with N_b buses interconnected by N_l transmission lines. Let $\mathcal{B} = \{1, 2, \dots, N_b\}$ denote the set of buses, and let $\mathcal{L} = \{1, 2, \dots, N_l\}$ denote the set of transmission lines. Each bus b could have some power generators, renewable supply and demand.

Assume that the RAC (Reliability Assessment Commitment) stage is conducted every T time slots. The purpose

of RAC is to ensure that future real-time dispatch decisions can always balance the supply and demand for all possible realizations of demand and renewable supply in the following T time slots. Below, we define both the characteristics of the demand and the power generating capabilities of the supply.

Demand: We first model the demand side. At each bus, there may exist many loads and many renewable energy plants. In this work, we assume that the renewable supply is cheaper than fossil-fuel generation, and will always be used to provide energy when available. As a result, the renewable supply can be viewed as negative demand. Thus, we only need to care about the net-demand, i.e., the total demand minus the total renewable supply, at each bus. Let $D_b(t)$ ($b \in \mathcal{B}, t = 1, 2, \dots, T$) be the net-demand at bus b and time t . Then, the entire net-demand time-sequence can be denoted as $D(1:T) = \{D_b(1:T), b \in \mathcal{B}\}$, where $D_b(1:T) = \{D_b(t), t = 1, 2, \dots, T\}$ is the net-demand time-sequence for each bus $b \in \mathcal{B}$.

Uncertainty: We now model the uncertainty due to renewable supply that the multi-stage decisions must deal with. Like [8][10], we assume that the net-demand sequence $D(1:T)$ may be any sequence in an uncertainty set \mathcal{D} , which is given as part of the problem formulation. In practice, this uncertainty set is typically constructed from forecasts performed before time 1 [8]. However, unlike the two-stage assumption in [8] that the entire sequence $D(1:T)$ is known precisely at the second stage, here at each time t , only the subsequence $D(1:t)$ is known, while the future subsequence $D(t+1:T)$ remains uncertain. While our proposed methodology can be applied to any form of uncertainty sets, for clarity we will use the following form for the rest of the paper. The uncertainty set \mathcal{D} consists of all $D(1:T)$ such that the following two constraints (1) and (2) are satisfied:

$$D_b^{\min}(t) \leq D_b(t) \leq D_b^{\max}(t), \quad (1)$$

$$\Delta_b^{\text{down}}(t_1, t_2) \leq D_b(t_1) - D_b(t_2) \leq \Delta_b^{\text{up}}(t_1, t_2), \quad (2)$$

where the parameters $D_b^{\min}(t)$ and $D_b^{\max}(t)$ denote the lower and upper limits, respectively, of the net-demand of bus b at time t , and the parameters $\Delta_b^{\text{up}}(t_1, t_2)$ and $\Delta_b^{\text{down}}(t_1, t_2)$ denote the maximum downward and upward speed of change for the net-demand of bus b . We note that although the uncertainty set \mathcal{D} is given for the entire time horizon, in a multi-stage setting the constraint (2) can be used to refine the remaining uncertainty at time t . Specifically, we introduce the following notation:

$$\mathcal{D}_{[t_1, t_2] | D(1:t)} = \{D(t_1:t_2) \mid \text{there exists } D'(1:T) \in \mathcal{D}, \text{ such that } D'(1:t) = D(1:t), D'(t_1:t_2) = D(t_1:t_2)\}, \quad (3)$$

which captures the remaining future uncertainty in the interval $[t_1, t_2]$, given the past net-demand trajectory $D(1:t)$. Note that this remaining uncertainty is usually smaller than the range of $D(t_1:t_2)$ in the original uncertainty set \mathcal{D} , since the past net-demand $D(1:t)$ has been fixed. If $t_1 = t_2$, we will simplify the notation as $\mathcal{D}_{[t_1, t_2]}$, which is just the original uncertainty set \mathcal{D} restricted to the time interval $[t_1, t_2]$.

Supply: We then model the supply side, i.e., the fossil-fueled generators in the power system. We use $\mathcal{G} = \{1, 2, \dots, N_g\}$ to denote the set of generators in the system,

and use $P_g(t)$ to denote the power level at each generator g and each time slot t . Denote

$$\mathbf{P}(1:T) = \{\mathbf{P}_g(1:T), g \in \mathcal{G}\} = \{P_g(t), g \in \mathcal{G}, t = 1, 2, \dots, T\}.$$

For a bus $b \in \mathcal{B}$, we use $\mathcal{G}_b \subseteq \mathcal{G}$ to denote the set of generators located at the bus b . Further, different generators have different power generating constraints. Specifically, for each generator $g \in \mathcal{G}$, $P_g(t)$ must be within the generators' capacity range $[P_g^{\text{minimum}}, P_g^{\text{maximum}}]$, i.e.,

$$P_g^{\text{minimum}} \leq P_g(t) \leq P_g^{\text{maximum}}, t = 1, 2, \dots, T, \quad (4)$$

Further, the power levels across time must obey the ramping constraints, i.e.,

$$|P_g(t) - P_g(t-1)| \leq R_g, t = 2, 3, \dots, T. \quad (5)$$

where R_g is the generator g 's ramping capability in one time slot (typically 5 minutes [1]).

Other Constraints: Given the demand $D(1:T)$, the power dispatch decision $\mathbf{P}(1:T)$ must satisfy both the demand-supply balance constraints and the transmission limit constraints. The demand-supply balance constraints require that the total power generation must be equal to the total net-demand (here, we ignore the transmission loss), i.e.,

$$\sum_{g \in \mathcal{G}} P_g(t) = \sum_{b \in \mathcal{B}} D_b(t), t = 1, 2, \dots, T. \quad (6)$$

To model the transmission limit constraints, we assume a simplified DC model [15]. Let $S = [S_{l,b}]$ be the shift factor of the power network, which is determined by the network topology and the reactance of the transmission lines. Then, the amount of power transmitting on line l at time slot t cannot exceed the power line l 's transmission limit, denoted by TL_l , i.e.,

$$\left| \sum_{b=1}^{N_b} S_{l,b} \left(D_b(t) - \sum_{g \in \mathcal{G}_b} P_g(t) \right) \right| \leq \text{TL}_l, \text{ for any } t, l. \quad (7)$$

A. Real Time Dispatch and RAC

As we discuss earlier, at each time t , the ISO must dispatch the generators so that the demand-supply balance and various constraints are met at time t . In practice, such dispatch decisions must respect *causality*, i.e., at each time t , the real-time decision can only be made based on the net-demand subsequence $D(1:t)$ already revealed and the uncertainty set $\mathcal{D}_{[t+1:T]|D(1:t)}$ (i.e., prediction) of the future subsequence $D(t+1:T)$, but cannot depend on the exact values of $D(t+1:T)$. Due to the causality requirement, it is possible that, if an incorrect decision was made at an earlier time, then at a future time, no dispatch decisions can meet all the constraints. As a result, the system may enter an unsafe state. Thus, it is imperative that the decisions at each time take future uncertainty into account, so that such unsafe scenario will never occur. Towards this end, we introduce the following concept.

Definition 1: Given an uncertainty set \mathcal{D} , we call $\pi(\mathcal{D})^2$ a causal real-time dispatch algorithm under \mathcal{D} , if at each time t ,

the dispatch decision $\mathbf{P}^{\pi(\mathcal{D})}(t) = \{P_g^{\pi(\mathcal{D})}(t), g \in \mathcal{G}\}$ produced by the algorithm $\pi(\mathcal{D})$ is a function of $D(1:t)$.

Definition 2: We say that a causal real-time dispatch algorithm $\pi(\mathcal{D})$ is *robust* for the uncertainty set \mathcal{D} , if and only if for all the demand sequence $D(1:T) \in \mathcal{D}$ and all time $t = 1, \dots, T$, the dispatch decision $\mathbf{P}^{\pi(\mathcal{D})}(t)$ produced by the algorithm $\pi(\mathcal{D})$ satisfies constraints (4)-(7).

The objectives of this work are then the following. First, at the RAC stage, given an uncertainty set \mathcal{D} , and a set of generators committed, we would like to know whether there exist causal real-time dispatch algorithms that are robust for the uncertainty set \mathcal{D} . Then, if the answer at the RAC stage is positive, we would like to find a causal real-time dispatch algorithm that is indeed robust for the uncertainty set \mathcal{D} . Note that unlike robust optimization [8], our formulation does not aim to minimize the worst-case future cost. As was discussed in the introduction and will be presented next, our solution produces "safe-dispatch sets" that allow the operator to balance both reliability and economy.

B. A Motivating Example for Multi-stage Decisions

Before we present our solution, we briefly contrast our problem formulation to typical two-stage methods in the literature for studying power-system operations under uncertainty, such as two-stage optimization [6][7][8][9]. A common feature of these two-stage methods is to assume that there exists a second stage where future uncertainty is fully revealed. In our setting, since the goal is to ensure that the physical constraints (4)-(7) are met at all times (see Definition 2), such a two-stage method would have corresponded to the following formulation:

Find a mapping from $D(1:T)$ to the dispatch decisions $\mathbf{P}(1:T)$ such that for all $D(1:T) \in \mathcal{D}$, the corresponding dispatch decision $\mathbf{P}(1:T)$ satisfies constraints (4)-(7) for all t .

However, such a mapping clearly violates the causality requirement in Definition 1. In practice, at any time $t < T$, only the subsequence $D(1:t)$ is known, while the future $D(t+1:T)$ remains unknown. Thus, the dispatch decisions computed by the above two-stage formulation cannot be executed in real time. More importantly, an RAC decision based on such a two-stage formulation would reach incorrect conclusions regarding system safety. Specifically, one may be tempted to declare the system to be safe under a given uncertainty set \mathcal{D} , whenever the above mapping in the two-stage formulation can be found for all $D(1:T) \in \mathcal{D}$. However, there may be no *causal* algorithms that are robust for the same uncertainty set \mathcal{D} ! The following simple example demonstrates this incorrect RAC decision³.

A Motivating Example: Suppose that there are two generators on a single bus. The first generator can operate between 0MW and 90MW, with ramping speed of +/-40MW per time-slot. The second generator can operate between 0MW and 10MW, with a ramping speed of +/-10MW per time-slot. Consider two possible net-demand sequences over 3 time-slots. The first sequence is (50,50,100)MW, and the second sequence is (50,50,0)MW. Under the two-stage formulation, for the first sequence, the mapped dispatch decision could be (50,50,90)MW for the first generator and (0,0,10)MW for

²Since \mathcal{D} is a part of the problem formulation, the algorithm π may thus behave differently for different uncertainty sets \mathcal{D} .

³A related but different example was also presented in [10].

the second generator. Similarly, for the second sequence, the mapped dispatch decision could be (40,40,0)MW for the first generator and (10,10,0)MW for the second generator. Thus, the two-stage formulation would declare such a setting as safe. However, it is easy to see that no causal algorithm can be robust for this level of uncertainty. Specifically, in order to reach 100MW in time-slot 3, the dispatch decision of the first generator at time-slot 2 must be greater than or equal to 50MW. On the other hand, in order to reach 0MW in time-slot 3, the dispatch decision of the first generator at time-slot 2 must be lower than or equal to 40MW. However, since the future demand at time-slot 3 is still unknown to the operator at time-slot 2, there exists no causal dispatch decision at time-slot 2 that can *simultaneously* meet the physical constraints for both possibilities.

This example illustrates the limitation of two-stage methods. In the following, we will present our solution for multi-stage decisions that correctly determines system reliability.

III. MAXIMALLY ROBUST ALGORITHM

Clearly, the two problems of RAC and real-time dispatch are tightly coupled. If the real-time dispatch algorithms are sub-optimal, more resources may need to be set aside at the RAC stage. Below, we first consider the real-time dispatch algorithm. Our goal is to design an “optimal” real-time dispatch algorithm regardless of the RAC decision. The notion of optimality is defined below.

Definition 3: Let $\Lambda = \{\mathcal{D} \mid \text{There exists a causal real-time dispatch algorithm } \pi_0 \text{ that is robust for the uncertainty set } \mathcal{D}\}$. A causal real-time dispatch algorithm π is said to be *maximally robust* if and only if $\pi(\mathcal{D})$ is robust for all the uncertainty sets $\mathcal{D} \in \Lambda$.

In other words, if any other algorithm is robust for an uncertainty set \mathcal{D} , then a maximally robust algorithm π must also be robust for \mathcal{D} . Thus, a maximally robust algorithm is the “most robust” among all algorithms. It turns out that this problem of determining dynamic control decisions so that the system states remain in a desired set of trajectories has been studied by dynamic programming (see section 4.6.2 in [11, p197]). In order to use the methodology of [11], we can treat the pair $D(1:t)$ and $\mathbf{P}(t)$ as the state of the system at the end of time t . Then, similar to the “target set” in [11, p197], we can introduce the notion of “safe-dispatch sets” as follows:

Definition 4: Given an uncertainty set \mathcal{D} , a demand history $D(1:t)$ and a power dispatch decision $\mathbf{P}(t)$, a causal real-time dispatch algorithm π is said to be *robust given $D(1:t)$ and $\mathbf{P}(t)$* , if and only if for any $D(t+1:T) \in \mathcal{D}_{[t+1,T] \mid D(1:t)}$, the algorithm π produces dispatch decisions $\{P_g^\pi(t_1), t_1 > t, g \in \mathcal{G}\}$ satisfying constraints (4)-(7) for all $t_1 > t$. (Note that (4)-(7) are defined for the entire range of t , but here $P_g^\pi(t_1)$ is only defined for $t_1 > t$.)

Definition 5: Given the demand history $D(1:t)$, the safe dispatch set $\mathcal{F}(D(1:t))$ at time t is defined as:

$$\mathcal{F}(D(1:t)) = \{\mathbf{P}(t) \mid \mathbf{P}(t) \text{ can balance the demand } D(t) \text{ subject to the constraints (4), (6) and (7), and there exists a causal algorithm } \pi \text{ that is robust given } D(1:t) \text{ and } \mathbf{P}(t)\}. \quad (8)$$

Intuitively, a maximally-robust algorithm simply needs to pick a dispatch decision at each time t from the safe-dispatch set $\mathcal{F}(D(1:t))$. As in [11, p197], this safe-dispatch set can be generated via backward induction. For ease of exposition, we use $A_t(D(t))$ to denote the set of all dispatch decisions that balance the net-demand $D(t)$ at time t subject to the generators’ capacity constraint and the transmission constraint:

$$A_t(D(t)) = \{\mathbf{P}(t) \mid \mathbf{P}(t), D(t) \text{ satisfy (4)(6)(7) at time } t\}. \quad (9)$$

Note that $A_t(D(t))$ does not include the ramping constraint (5). At $t = T$, the safe-dispatch set is simply given by (note that no ramping constraints need to be considered since T is the last time-slot):

$$\mathcal{F}(D(1:T)) = A_T(D(T)).$$

For all other $t < T$, suppose that $\mathcal{F}(D(1:t+1))$ is known for every possible $D(1:t+1)$. Let

$$f_t(A) = \{\mathbf{P}(t) \mid \text{There exists } \mathbf{P}(t+1) \in A, \text{ such that } |P_g(t+1) - P_g(t)| \leq R_g\}, \quad (10)$$

which can be interpreted as the set of dispatch decisions at time t that can ramp to one dispatch decision in A at time $t+1$. Then, the induction formula is given by

$$\mathcal{F}(D(1:t)) = \left(\bigcap_{D(t+1)} f_t(\mathcal{F}(D(1:t+1))) \right) \cap A_t(D(t)), \quad (11)$$

where the set intersection is taken over all $D(t+1) \in \mathcal{D}_{t+1 \mid D(1:t)}$ (recall that $D(1:t+1) = [D(1:t), D(t+1)]$). The detailed proof of the induction formula (11) is available in Appendix A. We now summarize the results that can be shown based on [11, p197]:

Proposition 6: Given the uncertainty set \mathcal{D} , there exists a causal and robust real-time dispatch algorithm if and only if $\mathcal{F}(D(1)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1 \triangleq \{D(1) \mid D(1:T) \in \mathcal{D}\}$.

Proof: See Appendix B. ■

Proposition 7: If $\mathcal{F}(D(1)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1$, then any algorithm in the following class is maximally robust:

Step 1: at time slot 1, pick an arbitrary dispatch decision $P(1) \in \mathcal{F}(D(1))$;

Step 2: at time slot $t > 1$, pick an arbitrary dispatch decision $\mathbf{P}(t) \in \mathcal{F}(D(1:t)) \cap C(\mathbf{P}(t-1))$, where

$$C(\mathbf{P}(t-1)) = \{\mathbf{P}(t) : |P_g(t) - P_g(t-1)| \leq R_g, g \in \mathcal{G}\} \quad (12)$$

is the set of dispatch decisions that can be reached at time t from the dispatch decision $\mathbf{P}(t-1)$ at time $t-1$.

Proof: See Appendix C. ■

According to Propositions 6 and 7, once we know how to calculate the safe-dispatch set $\mathcal{F}(D(1:t))$, both the RAC decision (Proposition 6) and the real-time dispatch decision (Proposition 7) are solved. However, in general the complexity of the backward induction (11) is high because there exist uncountably many demand sequences. Thus, the backward induction is useful only for theoretical analysis. In the next section, we will develop new algorithms for generating the safe-dispatch sets (or subsets) that are more computationally efficient.

IV. COMPUTATIONALLY-EFFICIENT ALGORITHMS

In this section, we study how to compute $\mathcal{F}(D(1:t))$ in polynomial time with respect to both T and N_b . Throughout this section, we will fix t and $D(1:t)$, and will derive $\mathcal{F}(D(1:t))$. We first study a simpler case where there are only one bus and two generators. In this case, we will show that $\mathcal{F}(D(1:t))$ can be exactly characterized by a polynomial number of linear constraints. We then study the general case with multiple buses and multiple generators, and proposed a new idea of *virtual demand splitting* (VDS), which can effectively compute a suitable subset of the safe dispatch set.

A. One Slow Generator + One Fast Generator

We first consider a one-bus two-generator case. We assume that the first generator is a slow generator, with time-varying generation limits $[P_{\text{slow}}^{\text{vmin}}(t), P_{\text{slow}}^{\text{vmax}}(t)]$, and time-varying up-ramping rate $R_{\text{slow}}^{\text{vup}}(t)$ and down-ramping rate $R_{\text{slow}}^{\text{vdown}}(t)$. Thus, the dispatched power level $P_{\text{slow}}^{\text{v}}(t)$ for this slow generator must satisfy

$$\begin{aligned} 0 &\leq P_{\text{slow}}^{\text{vmin}}(t) \leq P_{\text{slow}}^{\text{v}}(t) \leq P_{\text{slow}}^{\text{vmax}}(t), \\ -R_{\text{slow}}^{\text{vdown}}(t) &\leq P_{\text{slow}}^{\text{v}}(t+1) - P_{\text{slow}}^{\text{v}}(t) \leq R_{\text{slow}}^{\text{vup}}(t). \end{aligned}$$

The second generator is a fast generator. We assume that this fast generator can generate both negative and positive power in the range $[-r_{\text{fast}}^{\text{v-}}(t), r_{\text{fast}}^{\text{v+}}(t)]$, and there is no ramping constraint, i.e., this generator can ramp from any dispatch decision in $[-r_{\text{fast}}^{\text{v-}}(t), r_{\text{fast}}^{\text{v+}}(t)]$ to any dispatch decision in $[-r_{\text{fast}}^{\text{v-}}(t+1), r_{\text{fast}}^{\text{v+}}(t+1)]$. Then, the dispatched power level $P_{\text{fast}}^{\text{v}}(t)$ of the fast generator must satisfy

$$-r_{\text{fast}}^{\text{v-}}(t) \leq P_{\text{fast}}^{\text{v}}(t) \leq r_{\text{fast}}^{\text{v+}}(t).$$

Remark 1: It may seem unnatural to allow for negative power in the above constraints: this is to allow more flexibility in the generalization to the multi-bus multi-generator case in Section IV-B. Clearly, if one does not wish to allow negative power, the lower bound $r_{\text{fast}}^{\text{v-}}(t)$ can be set to 0. Hence, the above formulation is more general. Besides, this two-generator formulation also includes the single-generator case as a special case. Specifically, if we set both $r_{\text{fast}}^{\text{v-}}(t)$ and $r_{\text{fast}}^{\text{v+}}(t)$ to be 0, then this two-generator formulation reduces to a single-generator formulation.

Remark 2: Readers may notice that there is always a “v” in the superscript of the notations in this section. The reason is that in the future multi-bus multi-generator case, we will create virtual generator pairs, each of which corresponds to a pair of slow generator and fast generator as in this subsection. Thus, we add a superscript “v” (meaning “virtual”) to distinguish virtual generators from physical generators.

Since there is only one bus, $D(t)$ becomes a scalar. Thus, if we know the dispatch level $P_{\text{slow}}^{\text{v}}(t)$ of the slow generator, we can immediately obtain the dispatch level $P_{\text{fast}}^{\text{v}}(t)$ of the fast generator through $P_{\text{fast}}^{\text{v}}(t) = D(t) - P_{\text{slow}}^{\text{v}}(t)$. Hence, in the following discussion, we will only focus on the dispatch level $P_{\text{slow}}^{\text{v}}(t)$ of the slow generator. We first assume that $\mathcal{F}(D(1:t)) \neq \emptyset$, and derive some necessary conditions that the generator parameters ($P_{\text{slow}}^{\text{vmin}}(t)$, $P_{\text{slow}}^{\text{vmax}}(t)$, $R_{\text{slow}}^{\text{vdown}}(t)$, $R_{\text{slow}}^{\text{vup}}(t)$, $r_{\text{fast}}^{\text{v-}}(t)$, $r_{\text{fast}}^{\text{v+}}(t)$) need to satisfy. We then show that these conditions are also sufficient. As a result, we will obtain a close-form formula for $\mathcal{F}(D(1:t))$.

1) *Necessary Conditions for $\mathcal{F}(D(1:t)) \neq \emptyset$:* The first set of necessary conditions are quite obvious and they simply check whether the lower/upper limits of the slow generator are consistent with its ramping speed.

Lemma 8: (Parameter-checking condition) Given $D(1:t)$, if $\mathcal{F}(D(1:t)) \neq \emptyset$, then for any $t \leq t_0 \leq t_1$, the following conditions hold:

$$P_{\text{slow}}^{\text{vmin}}(t_0) - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s) \leq P_{\text{slow}}^{\text{vmax}}(t_1), \quad (13)$$

$$P_{\text{slow}}^{\text{vmax}}(t_0) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) \geq P_{\text{slow}}^{\text{vmin}}(t_1). \quad (14)$$

Clearly, if (13) is violated, then from any allowed power level at time t_0 (above $P_{\text{slow}}^{\text{vmin}}(t_0)$), the slow generator would have no way to ramp down to an allowed power level at t_1 (below $P_{\text{slow}}^{\text{vmax}}(t_1)$). Thus, $\mathcal{F}(D(1:t))$ would have been empty. The necessity of (14) is similar.

Even if conditions in (13) and (14) hold, the slow generator still may not be able to use all the power levels in $[P_{\text{slow}}^{\text{vmin}}(t_0), P_{\text{slow}}^{\text{vmax}}(t_0)]$. For instance, if $t_0 < t'$ and $P_{\text{slow}}^{\text{vmin}}(t_0) < P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)$, then the slow generator should not use any power level below $P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)$ at time t_0 . Otherwise, it will not be able to ramp up to any allowed power level (above $P_{\text{slow}}^{\text{vmin}}(t')$) at time t' . Similarly, if $t' < t_0$ and $P_{\text{slow}}^{\text{vmin}}(t_0) < P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)$, then from any allowed power level at t' (above $P_{\text{slow}}^{\text{vmin}}(t')$), the slow generator will never be able to reach below $P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)$ at time t_0 . Thus, we can define the “effective lower limit” of the slow generator at time $t_0 \geq t$ as

$$\begin{aligned} P_{\text{slow}}^{\text{eff-vmin}}(t_0) &= \max_{t_0 < t' \leq T} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s) \right\}, \\ &\max_{t \leq t' \leq t_0} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s) \right\}. \end{aligned} \quad (15)$$

Similarly, we can define the “effective upper limit” of the slow generator at time $t_0 \geq t$ as

$$\begin{aligned} P_{\text{slow}}^{\text{eff-vmax}}(t_0) &= \min_{t_0 < t' \leq T} \left\{ P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vdown}}(s) \right\}, \\ &\min_{t \leq t' \leq t_0} \left\{ P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vup}}(s) \right\}. \end{aligned} \quad (16)$$

Clearly, the power level of the slow generator at time t_0 should be within $[P_{\text{slow}}^{\text{eff-vmin}}(t_0), P_{\text{slow}}^{\text{eff-vmax}}(t_0)]$. The following necessary condition is then obvious.

Lemma 9: (Capacity condition) Given $D(1:t)$, if $\mathcal{F}(D(1:t)) \neq \emptyset$, then for any $t_0 \geq t$, the following conditions must hold:

$$\min_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \{D(t_0)\} \geq P_{\text{slow}}^{\text{eff-vmin}}(t_0) - r_{\text{fast}}^{\text{v-}}(t_0), \quad (17)$$

$$\max_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \{D(t_0)\} \leq P_{\text{slow}}^{\text{eff-vmax}}(t_0) + r_{\text{fast}}^{\text{v+}}(t_0). \quad (18)$$

In other words, no future demand can exceed the combined limits of the slow and fast generators.

While the above conditions are more obvious, the next condition is the key to capture the safety requirement in multi-stage decisions.

Lemma 10: (Load-following condition) Given $D(1:t)$, if $\mathcal{F}(D(1:t)) \neq \emptyset$, then for any $t \leq t_0 \leq \min\{t_1, t_2\}$, the following condition must hold:

$$\begin{aligned} & r_{\text{fast}}^{\text{v}+}(t_1) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) + r_{\text{fast}}^{\text{v}-}(t_2) + \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) \\ & \geq \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]|D(1:t)}} \left\{ \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \right. \\ & \quad \left. \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\} \right\}. \end{aligned} \quad (19)$$

Proof: Given any net-demand sequence $D(1:t_0) \in \mathcal{D}_{[1,t_0]|D(1:t)}$, consider the dispatch decision $(P_{\text{slow}}^{\text{v}}(t_0), P_{\text{fast}}^{\text{v}}(t_0))$ at time t_0 . We need to ensure that for any time $t_1 \geq t_0$, the maximum demand is reachable. (If not, the safe dispatch set $\mathcal{F}(D(1:t))$ would have been empty.) Thus, we have

$$P_{\text{slow}}^{\text{v}}(t_0) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) + r_{\text{fast}}^{\text{v}+}(t_1) \geq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\}.$$

Similarly, in order to reach the minimum demand at time $t_2 \geq t_0$, we must have

$$P_{\text{slow}}^{\text{v}}(t_0) - \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) - r_{\text{fast}}^{\text{v}-}(t_2) \leq \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\}.$$

Let

$$\begin{aligned} \gamma_{t_1}^{\min}(D(1:t_0)) & \triangleq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} \\ & - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) - r_{\text{fast}}^{\text{v}+}(t_1), \end{aligned} \quad (20)$$

and let

$$\begin{aligned} \gamma_{t_2}^{\max}(D(1:t_0)) & \triangleq \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\} \\ & + \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) + r_{\text{fast}}^{\text{v}-}(t_2). \end{aligned} \quad (21)$$

Then, we must have

$$\gamma_{t_1}^{\min}(D(1:t_0)) \leq P_{\text{slow}}^{\text{v}}(t_0) \leq \gamma_{t_2}^{\max}(D(1:t_0)), t_1, t_2 \geq t_0. \quad (22)$$

Substitute $\gamma_{t_1}^{\min}(D(1:t_0))$ and $\gamma_{t_2}^{\max}(D(1:t_0))$ by (20) and (21) in (22) and rearrange the obtained inequality. We can get

$$\begin{aligned} & r_{\text{fast}}^{\text{v}-}(t_2) + \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) + r_{\text{fast}}^{\text{v}+}(t_1) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) \\ & \geq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\}, \end{aligned}$$

The above inequality must hold for all possible net-demand subsequences $D(1:t_0) \in \mathcal{D}_{[1,t_0]|D(1:t)}$. We maximize the right-hand-side of the above inequality over all possible $D(1:t_0) \in \mathcal{D}_{[1,t_0]|D(1:t)}$, and the condition (19) then follows. \blacksquare

Remark 3: Note that the proof of Lemma 10 highlights a key difference between multi-stage and two-stage methods. Recall that, in the two-stage formulation in Section II-B, one assumes that the future demand is known at the second stage. As a result, for two-stage methods, one only need to check that, for *either* the maximum future demand at t_1 or the minimum future demand at t_2 , there exists a power level $P_{\text{slow}}^{\text{v}}(t_0)$ at time t_0 and a corresponding future dispatch trajectory that can balance the future demand. However, the two corresponding power-levels $P_{\text{slow}}^{\text{v}}(t_0)$ may differ depending on which future demand one wishes to check. In contrast, in the proof of Lemma 10, the operating point $P_{\text{slow}}^{\text{v}}(t_0)$ must be *simultaneously* capable of balancing both future demands. As readers may recall from the motivating example in Section II-B, this requirement leads to more restrictive conditions for system reliability, which is stated above in Lemma 10.

Remark 4: Since the uncertainty set \mathcal{D} only consists of linear constraints (1) and (2), it is easy to see that the terms “ $\min_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \{D(t_0)\}$ ”, “ $\max_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \{D(t_0)\}$ ” in (17)-(18), and the right-hand-side of (19), are all convex optimization problems, and thus can be computed efficiently.

2) *Sufficiency of Conditions (13)-(14),(17)-(19):* While the necessity of the above conditions (13)-(14),(17)-(19) are easy to follow, the next result is more surprising and it shows that these conditions are also sufficient for $\mathcal{F}(D(1:t)) \neq \emptyset$. Establishing this sufficiency is the first main contribution of our work.

Theorem 11: Given $D(1:t)$, if all the five conditions (13)-(14),(17)-(19) hold, then the safe dispatch set $\mathcal{F}(D(1:t))$ is not empty. Further, it can be explicitly expressed as follows:

$$\begin{aligned} \mathcal{F}(D(1:t)) & = \{(P_{\text{slow}}^{\text{v}}(t), P_{\text{fast}}^{\text{v}}(t)) | P_{\text{slow}}^{\text{v}}(t) \in h(D(1:t)), \\ & \quad P_{\text{slow}}^{\text{v}}(t) + P_{\text{fast}}^{\text{v}}(t) = D(t)\}. \end{aligned} \quad (23)$$

where $h(D(1:t))$ is an interval computed as follows

$$\begin{aligned} h(D(1:t)) & = \left[\max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t), \max_{t_1=t}^T \gamma_{t_1}^{\min}(D(1:t)) \right\}, \right. \\ & \quad \left. \min \left\{ P_{\text{slow}}^{\text{eff-vmax}}(t), \min_{t_1=t}^T \gamma_{t_1}^{\max}(D(1:t)) \right\} \right]. \end{aligned} \quad (24)$$

Note that the upper and lower limits in $h(D(1:t))$ are simply combinations of the limits in (22) and the effective limits $[P_{\text{slow}}^{\text{eff-vmin}}(t), P_{\text{slow}}^{\text{eff-vmax}}(t)]$. Thus, it naturally produces an outer bound on $\mathcal{F}(D(1:t))$. To show that $h(D(1:t))$ produces the exact form of $\mathcal{F}(D(1:t))$ as in (23), we will have to show that, for any $P(t) = (P_{\text{slow}}^{\text{v}}(t), P_{\text{fast}}^{\text{v}}(t))$ within the right hand side of (23), we can construct a causal real-time dispatch algorithm π (like in Prop. 7), such that this algorithm π is robust given $D(1:t)$ and $P(t)$. The detailed construction is in the proof of Theorem 11 (see Appendix D).

B. Multiple Buses + Multiple Generators

Part of the reason why the above two-generator case is easier is because the allowed dispatch level of the slow generator

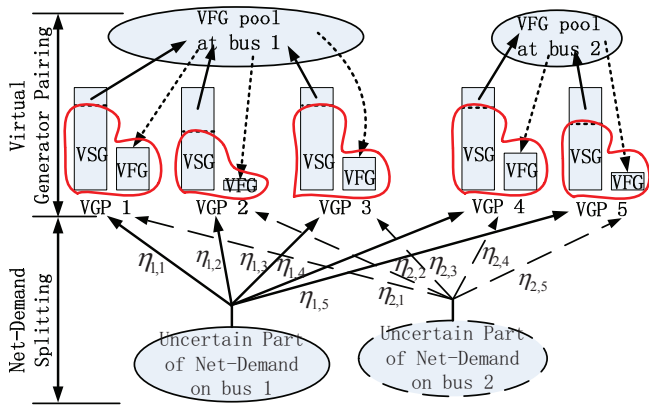


Fig. 1. Illustration of the VDS approach

can be captured by a one-dimensional interval. Unfortunately, this is no longer true when we move to the general case of multiple generators and multiple buses, which becomes more difficult. In this subsection, we focus on obtaining a subset $\mathcal{F}^{\text{VDS}}(D(1:t)) \subset \mathcal{F}(D(1:t))$ for the general case. Then, if we can show that $\mathcal{F}^{\text{VDS}}(D(1:t)) \neq \emptyset$, we will immediately have $\mathcal{F}(D(1:t)) \neq \emptyset$. Our basic idea is *demand splitting*, i.e., fractions of the future net-demand uncertainty are sent to the generators according to pre-computed *splitting factors*. However, due to the physical constraints of the generators, assigning a splitting factor for each physical generator will lead to severely-reduced safe-dispatch sets, an example of which is shown in Appendix E. This observation thus motivates us to consider the following idea of *virtual demand splitting*, where a splitting factor is assigned to each pair of virtual slow generator (VSG) and virtual fast generator (VFG). The corresponding fraction of net-demand uncertainty is then sent to such a virtual generator pair (VGP). For each VGP, we can then use the two-generator characterization in Section IV-A. This pairing of virtual generators is performed using the concept of *VFG pool* defined below. Before we describe the details, we emphasize that such pairing relationships and the corresponding splitting factors computed by the procedure below is only used to determine a subset $\mathcal{F}^{\text{VDS}}(D(1:t))$ of the safe-dispatch set $\mathcal{F}(D(1:t))$. They are *not* used for the actual dispatch of the generators. Rather, once $\mathcal{F}^{\text{VDS}}(D(1:t))$ is known, the actual dispatch decisions can be done as described in Section IV-B4 to balance both reliability and economy.

1) *Creating VGPs (Virtual Generator Pairs)*: This part involves the following steps and parameters (see also Fig. 1). (i) There is a VFG (Virtual Fast Generator) pool on each bus b . (ii) Each generator $g \in \mathcal{G}_b$ on bus b contributes $R_{\text{fast},g}^{\text{V}+}(t_0), t_0 \geq t$ to the upward direction of the VFG pool on the bus, and contributes $R_{\text{fast},g}^{\text{V}-}(t_0), t_0 \geq t$ to the downward direction (see the upper solid arrows in Fig. 1). Thus, the overall VFG pool is able to ramp from any point in $[-\sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}-}(t_0), \sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}+}(t_0)]$ to any point in $[-\sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}-}(t_0+1), \sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}+}(t_0+1)]$ in one time-slot. (iii) Each generator g reduces its ramping speed and power limits to $R_{\text{slow},g}^{\text{Vup}}, R_{\text{slow},g}^{\text{Vdown}}, P_{\text{slow},g}^{\text{Vmax}}$ and $P_{\text{slow},g}^{\text{Vmin}}$ according to how much it contributes to the VFG pool. Specifically,

$$R_{\text{slow},g}^{\text{Vup}}(t_0) = R_g - R_{\text{fast},g}^{\text{V}+}(t_0+1) - R_{\text{fast},g}^{\text{V}-}(t_0) \geq 0, \quad (25)$$

$$R_{\text{slow},g}^{\text{Vdown}}(t_0) = R_g - R_{\text{fast},g}^{\text{V}-}(t_0+1) - R_{\text{fast},g}^{\text{V}+}(t_0) \geq 0, \quad (26)$$

$$P_{\text{slow},g}^{\text{Vmax}}(t_0) = P_g^{\text{maximum}} - R_{\text{fast},g}^{\text{V}+}(t_0), \quad (27)$$

$$P_{\text{slow},g}^{\text{Vmin}}(t_0) = P_g^{\text{minimum}} + R_{\text{fast},g}^{\text{V}-}(t_0). \quad (28)$$

These parameters define a remaining generator with reduced capabilities, which we refer to as the VSG (Virtual Slow Generator) g . (iv) Each VSG g is then paired with a VFG with range $[-r_{\text{fast},g}^{\text{V}-}(t_0), r_{\text{fast},g}^{\text{V}+}(t_0)]$. They combined form a VGP (virtual generator pair). All the VFGs on a bus come from the same VFG pool on the same bus (see the upper dashed arrows in Fig. 1). Hence, the total range of the VFGs on a bus cannot exceed the range of the VFG pool on the same bus, i.e.,

$$\sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}+}(t_0) \geq \sum_{g \in \mathcal{G}_b} r_{\text{fast},g}^{\text{V}+}(t_0), \quad (29)$$

$$\sum_{g \in \mathcal{G}_b} R_{\text{fast},g}^{\text{V}-}(t_0) \geq \sum_{g \in \mathcal{G}_b} r_{\text{fast},g}^{\text{V}-}(t_0). \quad (30)$$

2) *Demand Splitting for VGP*: For every $t_0 \geq t$, we divide the net-demand $D_b(t_0)$ into two parts, i.e., $D_b(t_0) = D_b^{\text{main}}(t_0) + (D_b(t_0) - D_b^{\text{main}}(t_0))$, where $D_b^{\text{main}}(t_0) = (\max\{D_b(t_0)\} + \min\{D_b(t_0)\})/2$ is called the main part of $D_b(t_0)$, and $(D_b(t_0) - D_b^{\text{main}}(t_0))$ is called the uncertain part.

In order to characterize $\mathcal{F}^{\text{VDS}}(D(1:t))$, we assume that the following dispatch decisions will be carried out in the future. Specifically, we dispatch⁴ the two parts of net-demand separately. For the main part, we dispatch $P_{\text{VGP},g}^{\text{main}}(t_0)$ amount of power to each VGP g such that

$$\sum_{b=1}^{N_b} D_b^{\text{main}}(t_0) = \sum_{g=1}^{N_g} P_{\text{VGP},g}^{\text{main}}(t_0).$$

For the uncertain part, we introduce the concept of the splitting factor $\eta = [\eta_{b,g}, b \in \mathcal{B}, g \in \mathcal{G}]$. For each bus b and each VGP g , $\eta_{b,g}$ is the fraction of the uncertain part of $D_b(t_0)$ to be dispatched to the VGP g (see the lower arrows in Fig. 1). Clearly, we need

$$\sum_{g \in \mathcal{G}} \eta_{b,g} = 1, \text{ for all } b \in \mathcal{B}.$$

Thus, each VGP will be allocated $\sum_{b \in \mathcal{B}} \eta_{b,g} (D_b(t_0) - D_b^{\text{main}}(t_0))$ amount of demand from the uncertain part of $D_b(t_0)$. Hence, the total demand $D_g(t_0)$ allocated to the VGP g at time slot t is

$$D_g(t_0) = P_{\text{VGP},g}^{\text{main}}(t_0) + \sum_{b \in \mathcal{B}} \eta_{b,g} (D_b(t_0) - D_b^{\text{main}}(t_0)). \quad (31)$$

Finally, the quantities $\{D_g(t_0), g \in \mathcal{G}, t_0 = t, \dots, T\}$ need to satisfy the following transmission limits (TL):

$$\max_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \sum_{b=1}^{N_b} S_{l,b} \left(D_b(t_0) - \sum_{g \in \mathcal{G}_b} D_g(t_0) \right) \leq \text{TL}_l \quad (32)$$

⁴As we discussed earlier, although we use the word “dispatch”, the corresponding equations are only used to compute the subset $\mathcal{F}^{\text{VDS}}(D(1:t))$. They are not used to compute the actual dispatch decision in real time.

$$\min_{D(t_0) \in \mathcal{D}_{t_0|D(1:t)}} \sum_{b=1}^{N_b} S_{l,b} \left(D_b(t_0) - \sum_{g \in \mathcal{G}_b} D_g(t_0) \right) \geq -\text{TL}_l \quad (33)$$

Remark 5: Note that both (32) and (33) are convex constraints. Take (32) for example. For each value of $D(t_0)$, $\sum_{b=1}^{N_b} S_{l,b}(D_b(t_0) - \sum_{g \in \mathcal{G}_b} D_g(t_0))$ is a linear function of $\bar{P}_g^v(t_0)$ and $\eta_{b,g}$. Then, the left-hand-side is a maximum of linear functions, and hence is convex [16]. Similarly, (33) is also convex.

3) *Deriving $\mathcal{F}^{\text{VDS}}(D(1:t))$:* We are now ready to give the detailed definition of $\mathcal{F}^{\text{VDS}}(D(1:t))$.

The set of variables $Z(t) = \{\eta_{b,g}, R_{\text{fast},g}^{v-}(t_0), R_{\text{fast},g}^{v+}(t_0), r_{\text{fast},g}^{v-}(t_0), r_{\text{fast},g}^{v+}(t_0), P_{\text{VGP},g}^{\text{main}}(t_0), t_0 \geq t, b \in \mathcal{B}, g \in \mathcal{G}\}$ are under our control. Each $Z(t)$ gives a parameterization of VGP g for each g and its demand uncertainty. We can then apply Theorem 11 to obtain a set of constraints under which the safe dispatch set for each VGP g , denoted by $\mathcal{F}_g^{Z(t)}(D(1:t))$, is not empty. Note that any set of safe dispatch decisions $(P_{\text{slow},g}^v(t), P_{\text{fast},g}^v(t)) \in \mathcal{F}_g^{Z(t)}(D(1:t))$ for all VGPs g , can be mapped to dispatch decisions $\mathbf{P}(t)$ on real generators that satisfy the following constraints:

$$-R_{\text{fast},g}^{v-}(t) \leq P_g(t) - P_{\text{slow},g}^v(t) \leq R_{\text{fast},g}^{v+}(t), \quad (34)$$

$$\sum_{g \in \mathcal{G}_b} P_g(t) = \sum_{g \in \mathcal{G}_b} (P_{\text{slow},g}^v(t) + P_{\text{fast},g}^v(t)). \quad (35)$$

In (34), $P_g(t) - P_{\text{slow},g}^v(t)$ is the amount of power generated by the VFG contributed from the physical generator g , and thus must be within the range $[-R_{\text{fast},g}^{v-}(t), R_{\text{fast},g}^{v+}(t)]$. In (35), the total power generated by the VGPs at one bus must be equal to the total power generated by the physical generators at the same bus. Finally, any dispatch decision $\mathbf{P}(t) = \{P_g(t), g \in \mathcal{G}\}$ obtained in this way will also balance the net-demand. Hence, we define $\mathcal{F}^{\text{VDS}}(D(1:t))$ as follows:

$$\begin{aligned} \mathcal{F}^{\text{VDS}}(D(1:t)) = \{ & \mathbf{P}(t) | \text{There exists } Z(t) \text{ satisfying} \\ & \text{all admissible constraints in Sec. IV-B1 and IV-B2} \\ & \text{including (25)-(30) and (32)-(33), and there exists} \\ & (P_{\text{slow},g}^v(t), P_{\text{fast},g}^v(t)) \in \mathcal{F}_g^{Z(t)}(D(1:t)) \neq \emptyset, g \in \mathcal{G}, \\ & \text{satisfying (34) and (35)} \} \end{aligned} \quad (36)$$

Clearly, $\mathcal{F}^{\text{VDS}}(D(1:t))$ is a subset of $\mathcal{F}(D(1:t))$. Further, it consists of a polynomial number of convex constraints.

Remark 6: We note that our idea of demand splitting shares some similarity to the choice of affine policies in the multi-stage robust-optimization approach in [10]. However, there are two key differences. First, we assign a splitting factor for each pair of VSG and VFG. This is possible because we have already derived in Section IV-A the precise conditions for the safe-dispatch sets given such a pair of generators. In contrast, [10] assigns a splitting factor for each physical generator. Note that the latter can be viewed as a special case of the former by setting all the VFGs to zero. As we argue earlier (and see details in Appendix E), the additional flexibility by using VGPs will likely enlarge the obtained subset $\mathcal{F}^{\text{VDS}}(D(1:t))$. Second, unlike [10], we do not use the splitting factors for the actual real-time dispatch. Rather, they are only used for computing $\mathcal{F}^{\text{VDS}}(D(1:t))$. Once $\mathcal{F}^{\text{VDS}}(D(1:t))$ is

known, in real-time one can choose any operating point within this subset, e.g., the one that is most economic. The details are provided next.

4) *Balancing Reliability and Economy:* Once we obtain a subset $\mathcal{F}^{\text{VDS}}(D(1:t))$ of the safe-dispatch set, in the RAC stage we can simply verify system reliability by checking $\mathcal{F}^{\text{VDS}}(D(1:t)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1 \triangleq \{D(1) | D(1:T) \in \mathcal{D}\}$. Further, in real-time dispatch we can combine with economic dispatch to obtain the following VDS-ED (Virtual Demand Splitting-Economic Dispatch) algorithm that balances both reliability and economy.

- 1 At time slot 1, pick the dispatch decision $\mathbf{P}(1) \in \mathcal{F}^{\text{VDS}}(D(1))$ that minimizes the generation cost $\sum_{g \in \mathcal{G}} \text{Cost}_g(P_g(1))$ at time 1, where $\text{Cost}_g(\cdot)$ is the cost function for generator g .
- 2 At time slot $t > 1$, pick the dispatch decision $P(t) \in \mathcal{F}^{\text{VDS}}(D(1:t)) \cap C(\mathbf{P}(t-1))$ that minimizes the generation cost $\sum_{g \in \mathcal{G}} \text{Cost}_g(P_g(t))$ at time t , where $C(\mathbf{P}(t-1))$ is given by (12).

Algorithm 1: VDS-ED Algorithms

The robustness of the VDS-ED algorithm can be shown similar to Prop. 7 (see Appendix F for the detailed proof). To see how the VDS-ED algorithm allows us to balance both reliability and economy, suppose that the standard real-time ED (Economic Dispatch), which minimizes the generation cost at each time t without the constraints from the set $\mathcal{F}^{\text{VDS}}(D(1:t))$, produces a decision that falls within the set $\mathcal{F}^{\text{VDS}}(D(1:t))$. Then, the VDS-ED algorithm will obviously follow the same dispatch decision as the ED algorithm. On the other hand, if the decision of the ED algorithm exceeds the set $\mathcal{F}^{\text{VDS}}(D(1:t))$, the VDS-ED will modify the dispatch decision to be within the set. In this way, the VDS-ED algorithm can be viewed as the ‘‘robustified’’ version of the standard ED algorithm [17]: it only intervenes when necessary. As a result, the VDS-ED algorithm achieves robustness in the worst case, without sacrificing economy in the average case. We emphasize that this flexibility is new to the multi-stage robust optimization work in [10], which requires that real-time decision must follow the affine policy that are tailored to the worst case. Finally, even though the set $\mathcal{F}^{\text{VDS}}(D(1:t))$ looks quite complicated, it is convex and can be written as a polynomial number of linear or convex constraints. Therefore, the optimization problems in VDS-ED can be effectively solved.

V. SIMULATION

In this section, we present preliminary simulation results based on the 4-bus system in Fig. 2(a). Although this system is undoubtedly small, it allows us to illustrate the key advantage of our proposed approach compared to the two-stage methods and the standard economic dispatch. (More extensive simulation results will be reported in our future work.) This 4-bus system has 4 transmission lines, each of which has a transmission limit of 2500 MW. There is load at buses 2 and 3, and wind supply at bus 3. The demand data and the renewable energy data are both borrowed from Elia [18], Belgium’s electricity transmission system operator. Specifically, we split the load data from 6am to 9am on 01/01/2015 evenly into two

parts, and feed them into buses 2 and 3 (see Fig. 2(b)). The wind data is also from 6am to 9am on 01/01/2015 (see Fig. 2(c)), and we simply feed it to bus 3. Note that the load data is more predictable than the wind data. Hence, for simplicity, we assume that the exact values of the load are known at the RAC stage, and thus the uncertainty all comes from the wind energy. The uncertainty set of the wind is modeled according to (1) and (2). Specifically, the upper bound and the lower bound are shown in Fig. 2(c), and the maximum variation $\Delta_w^{\text{up}}(\Delta t) = \Delta_w^{\text{up}}(t, t + \Delta t)$ and $\Delta_w^{\text{down}}(\Delta t) = \Delta_w^{\text{down}}(t, t + \Delta t)$ of wind (here we assume that the maximum variation only depends on the time difference) is shown in Fig. 2(d).

This system has fossil-fueled generators at buses 1 and 4. We use 4 types of fossil-fueled generators in our simulation, which are listed in Table I. In our 4-bus system, there are 2 type-A generators, 3 type-B generators, and 7 type-D generators at bus 1, and there are 2 type-B generators, 5 type-C generators, and 6 type-D generators at bus 4.

TABLE I. LISTS OF FOSSIL-FUELED GENERATORS

Type	Generator Limits	Ramping Rate	Price
A	600-1200MW	30MW/min	15\$/MWh
B	420-600MW	4MW/min	10\$/MWh
C	300-900MW	30MW/min	15\$/MWh
D	0-200MW	100MW/min	30\$/MWh

We use the VDS approach at the RAC stage to verify power system safety under different penetration levels of wind energy (i.e., by checking whether $\mathcal{F}^{\text{VDS}}(D(1:t)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1 = \{D(1)|D(1:T) \in \mathcal{D}\}$). We then compare the real-time dispatch of both the VDS-ED algorithm and the standard ED algorithm. (The ED algorithm, which is commonly used by the ISOs [1], picks a $P(t)$ satisfying (4)-(7) that minimizes the total power-generation cost at each time t .) Specifically, we scale up the wind energy (together with its bounds and maximum variation) by a scaling factor ranging from 1 to 2.5, and summarize the results in Table II. We can see that, as long as the system is safe under VDS in the RAC stage (the first row), the VDS-ED algorithm can also ensure system safety in real time (the second row). In contrast, if the standard ED algorithm is used for real-time dispatch, the system can be unsafe at high wind penetration levels (the third row), even though it could be safe when VDS-ED is used.

TABLE II. SAFETY CHECK AND REAL TIME DISPATCH ("1" MEANS SAFE)

scaling factor	1	1.4	1.8	1.9	2.1	2.2	2.3	2.6
VDS-RAC	1	1	1	1	1	1	0	0
VDS-ED	1	1	1	1	1	1	0	0
ED	1	1	1	0	0	0	0	0
Two-Stage	1	1	1	1	1	1	1	1

We also compare with the two-stage formulation in Section II-B for checking system safety. As readers can see from the fourth row of Table II, even when the scaling factor is as large as 2.6, the two-stage formulation still declares that the system is safe. However, as we discussed in Section II-B, in practice one cannot use the dispatch decisions produced by such a two-stage formulation. If we just used standard ED algorithm, the system would be unsafe once the scaling factor is beyond 1.8. Even if we used the VDS-ED algorithm, the system would still be unsafe for scaling factors above 2.2. Between 2.2 and 2.6,

we do not know whether there exist causal real-time dispatch algorithms that can be robust at the corresponding levels of uncertainty (because the VDS-ED algorithm is based on a subset of the true safe-dispatch set). Nonetheless, this example illustrates that the two-stage formulations alone can easily lead to incorrect conclusions on system safety. In contrast, our proposed VDS approach produces consistent and correct conclusions for both the real-time and RAC stages.

We also compare the total generation cost (i.e., the summation of the cost at all times) between the VDS-ED algorithm and the standard ED algorithm. We can see from Table III that, when the ED algorithm is safe, which is the case when the scaling factor is 1, 1.4 and 1.8, our VDS-ED achieves the same cost. The reason is that, if the ED decision is within the safe dispatch set computed from the VDS approach, our VDS-ED algorithm will follow the same dispatch decision. Only when the ED decision is outside the safety set, our VDS-ED algorithm will modify the ED decision to another decision within the safety set, which occurs when the scaling factor is 1.9. (A more detailed analysis based on simulation data is provided in Appendix G.) Based on this simulation, we can see that the VDS-ED algorithm achieves a higher level of robustness than the standard ED algorithm, without sacrificing much economy.

TABLE III. TOTAL COST COMPARISON

scaling factor	1	1.4	1.8	1.9
VDS-ED	234320\$	218255\$	204928\$	201760\$
ED	234320\$	218255\$	204928\$	Inf

VI. CONCLUSION

We study online multi-stage decisions to ensure grid safety under high renewable uncertainty. Using the concept of safe dispatch sets, we first construct a class of maximally robust algorithms for real-time dispatch. Further, the non-emptiness of the safe dispatch sets also provides a natural condition for verifying grid safety at the RAC stage. Unfortunately, computing such safe-dispatch sets is in general very difficult. We then develop efficient methods to exactly characterize the safe-dispatch set for the two-generator case and compute a suitable subset of the safe-dispatch set in the general case. Our simulation results show that the resulting VDS-ED algorithm outperforms the standard real-time economic dispatch algorithm in terms of robustness, without sacrificing economy. For future work, we will perform more extensive experiments for larger settings, such as the IEEE power system test cases [19]. Further, we will study how to integrate the safety condition at the RAC stage with unit-commitment to determine the most suitable set of resources that should be set aside in advance.

REFERENCES

- [1] <https://www.misoenergy.org/>.
- [2] <http://www.caiso.com/>.
- [3] <http://www.ercot.com/>.
- [4] D. Bienstock and A. Verma, "The N - k Problem in Power Grids: New Models, Formulations and Numerical Experiments," *SIAM Journal on Optimization*, vol. 20, pp. 2352–2380, 2010.
- [5] Q. Wang, J.-P. Watson, and Y. Guan, "Two-Stage Robust Optimization for N - k Contingency-Constrained Unit Commitment," *IEEE Transactions on Power Systems*, August 2013.

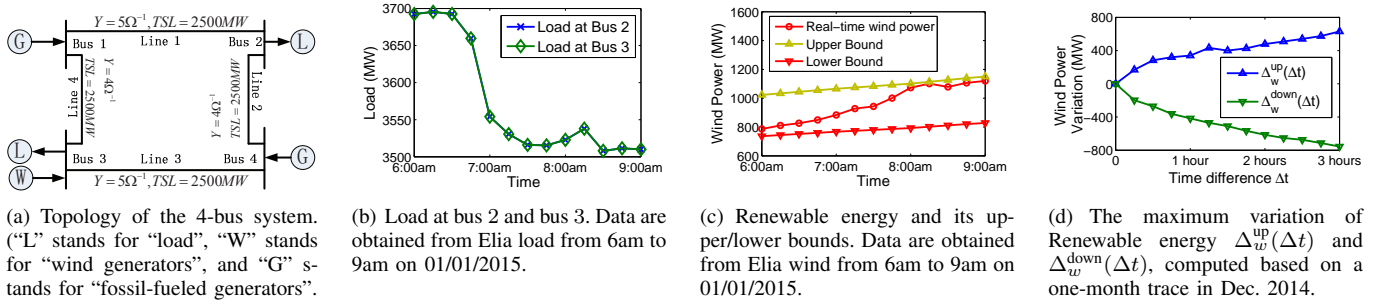


Fig. 2. Simulation data: (a) topology; (b) load; (c) wind and its range; (d) maximum variation of wind.

- [6] Q. P. Zheng, J. Wang, P. M. Pardalos, and Y. Guan, "A decomposition approach to the two-stage stochastic unit commitment problem," *Annals of Operations Research*, November 2013.
- [7] P. P. Varaiya, F. F. Wu, and J. W. Bialek, "Smart Operation of Smart Grid: Risk-Limiting Dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40–57, 2011.
- [8] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. on Power Systems*, vol. 28, no. 1, 2013.
- [9] C. Zhao and Y. Guan, "Unified stochastic and robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3353–3361, 2013.
- [10] A. Lorca, A. Sun, E. Litvinov, and T. Zheng, "Multistage Adaptive Robust Optimization for the Unit Commitment Problem," *Available at <http://www2.isye.gatech.edu/%7Exsun84/publications/>*, 2014.
- [11] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific Belmont, MA, 2007, vol. 1.
- [12] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource Allocation and Cross-Layer Control in Wireless Networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.
- [13] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of Robust Optimization," *SIAM review*, vol. 53, no. 3, 2011.
- [14] A. Borodin and R. El-Yaniv, *Online Computation and Competitive Analysis*. Cambridge University Press, 2005.
- [15] A. J. Wood and B. F. Wollenberg, *Power generation operation and control*. Wiley-Interscience, 1996.
- [16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [17] S. Zhao, X. Lin, and M. Chen, "Peak-Minimizing Online EV Charging: Price of Uncertainty and Algorithm Robustification," in *IEEE INFO-COM*, Hong Kong, China, April 2015.
- [18] <http://www.elia.be/en/grid-data/>.
- [19] <https://www.ee.washington.edu/research/pstca/>.

APPENDIX A PROOF OF EQUATION (11)

Proof: In order to prove Equation (11), we first show that

$$\mathcal{F}(D(1:t)) \subseteq \left(\bigcap_{D(t+1)} f_t(\mathcal{F}(D(1:t+1))) \right) \cap A_t(D(t)). \quad (37)$$

Consider an arbitrary dispatch decision $\mathbf{P}(t) \in \mathcal{F}(D(1:t))$. Then, there exists an algorithm π that is robust given $D(1:t)$ and $\mathbf{P}(t)$. This indicates that for any $D(t+1) \in \mathcal{D}_{t+1|D(1:t)}$, if the dispatch decision at time $t+1$ is $P^\pi(t+1)$, the algorithm π will also be robust given $D(1:t+1)$ and $P^\pi(t+1)$. Hence, $P^\pi(t+1) \in \mathcal{F}(D(1:t+1))$. According to the definition (10)

of the function $f_t(\cdot)$, we then have

$$\mathbf{P}(t) \in f_t(\mathcal{F}(D(1:t+1))) \quad (38)$$

Note that (38) holds for all $D(t+1) \in \mathcal{D}_{t+1|D(1:t)}$, and $\mathbf{P}(t) \in A_t(D(t))$ according to (8). Therefore,

$$\mathbf{P}(t) \in \left(\bigcap_{D(t+1)} f_t(\mathcal{F}(D(1:t+1))) \right) \cap A_t(D(t)).$$

Thus, (37) holds.

We then show that

$$\mathcal{F}(D(1:t)) \supseteq \left(\bigcap_{D(t+1)} f_t(\mathcal{F}(D(1:t+1))) \right) \cap A_t(D(t)). \quad (39)$$

We only need to show that for any

$$\mathbf{P}(t) \in \left(\bigcap_{D(t+1)} f_t(\mathcal{F}(D(1:t+1))) \right) \cap A_t(D(t)),$$

there exists a causal algorithm π that is robust given $D(1:t)$ and $\mathbf{P}(t)$.

Note that $\mathbf{P}(t) \in f_t(\mathcal{F}(D(1:t+1)))$ for all $D(t+1) \in \mathcal{D}_{t+1|D(1:t)}$. Hence, for any $D(t+1) \in \mathcal{D}_{t+1|D(1:t)}$, there exists a dispatch decision $P^{D(t+1)}(t+1)$ such that $|P_g^{D(t+1)}(t+1) - P_g(t)| \leq R_g$ for any $g \in \mathcal{G}$, and there exists an algorithm, denoted by $\pi_{D(t+1)}$, that is robust given $D(1:t+1)$ and $P^{D(t+1)}(t+1)$. We then construct an algorithm π as follows:

- At time $t+1$, based on $D(t+1)$, the algorithm π set its dispatch as $P^{D(t+1)}(t+1)$;
- After time $t+1$, the algorithm π uses the same dispatch as the algorithm $\pi_{D(t+1)}$.

It is easy to check that π is robust given $D(1:t)$ and $\mathbf{P}(t)$. Note that $\mathbf{P}(t)$ is also in $A_t(D(t))$. Then, according to Definition 5, we must have $\mathbf{P}(t) \in \mathcal{F}(D(1:t))$. Equation (39) then follows.

Combining (37) and (39), we complete the proof. \blacksquare

APPENDIX B
PROOF OF PROPOSITION 6

Proof: Necessity: Let π be a causal real-time dispatch algorithm that is robust for the uncertainty set \mathcal{D} . Let $P^\pi(1)$ be the dispatch decision at time-slot 1 under the algorithm π that balances the demand $D(1) \in \mathcal{D}_1$. Obviously, this algorithm π is robust given $D(1)$ and $P^\pi(1)$. Thus, $P^\pi(1) \in \mathcal{F}(D(1))$, which indicates that $\mathcal{F}(D(1))$ is not empty for all $D(1) \in \mathcal{D}_1$.

Sufficiency: Given $\mathcal{F}(D(1)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1$, we only need to construct a causal algorithm π that is robust for the uncertainty set \mathcal{D} . Specifically,

(i) at time 1, based on the demand $D(1)$, this algorithm π picks an arbitrary dispatch decision $P(D(1))$ from $\mathcal{F}(D(1))$.

Based on Definition 5, we know that there must exist a causal real-time dispatch algorithm $\pi^{D(1)}$ that is robust given $D(1)$ and $P(D(1))$. Then,

(ii) after time 1, the algorithm π simply follows the algorithm $\pi^{D(1)}$.

Based on the above construction of the algorithm π , it is easy to see that the algorithm π is robust for the uncertainty set \mathcal{D} . ■

APPENDIX C
PROOF OF PROPOSITION 7

In order to prove Proposition 7, we only need to show that any algorithm defined in Proposition 7 can produce safe dispatch decisions for any demand sequence $D(1:T) \in \mathcal{D}$. Specifically, we need to show that $\mathcal{F}(D(1))$ and $\mathcal{F}(D(1:t)) \cap C(\mathbf{P}(t-1)), t = 2, \dots, T$ are all nonempty.

Proof: We have assumed that $\mathcal{F}(D(1))$ is not empty, then the algorithm defined in Proposition 7 can give a valid dispatch decision at time 1. Next, we prove by induction. Assume that the algorithm defined in Proposition 7 can pick a valid dispatch decision at time $t-1$. We will show $\mathcal{F}(D(1:t)) \cap C(\mathbf{P}(t-1)) \neq \emptyset$.

Note that $\mathbf{P}(t-1)$ must be in $\mathcal{F}(D(1:t-1))$. Based on the induction formula (11), it is easy to see that $\mathbf{P}(t-1) \in f_t(\mathcal{F}(D(1:t)))$. Then, by definition (10) of $f_t(\cdot)$, there must exist $\mathbf{P}(t) \in \mathcal{F}(D(1:t))$ such that $|P_g(t) - P_g(t-1)| \leq R_g$. Hence, $\mathbf{P}(t) \in \mathcal{F}(D(1:t)) \cap C(\mathbf{P}(t-1))$, and thus $\mathcal{F}(D(1:t)) \cap C(\mathbf{P}(t-1))$ is nonempty.

Based on Proposition 6, whenever there exists a causal and robust real-time dispatch algorithm for the uncertainty set \mathcal{D} , $\mathcal{F}(D(1))$ will be nonempty for all $D(1) \in \mathcal{D}_1$. The above analysis implies that whenever $\mathcal{F}(D(1)) \neq \emptyset$ for all $D(1) \in \mathcal{D}_1$, any algorithm defined in Proposition 7 will be robust for the uncertainty set \mathcal{D} . Combining Proposition 6 with the above analysis, we can then reach the conclusion that any algorithm defined in Proposition 7 is maximally robust. ■

APPENDIX D
PROOF OF THEOREM 11

We use I to denote the set on the right-hand-side of (23). The key in the proof is to construct a causal real-time dispatch algorithm π that is robust given $D(1:t)$ and starting from any

$\mathbf{P}(t) = (P_{\text{slow}}^v(t), P_{\text{fast}}^v(t))$ within the set I . (Note that any $\mathbf{P}(t) \in I$ can balance the demand $D(t)$ at time t .)

Consider the following causal real-time dispatch algorithm π :

Algorithm π : At any time slot $t_0 > t$, pick an arbitrary dispatch decision $P_{\text{slow}}^v(t_0) \in h(D(1:t_0)) \cap C(P_{\text{slow}}^v(t_0-1))$, and set $P_{\text{fast}}^v(t_0) = D(t_0) - P_{\text{slow}}^v(t_0)$. Here, $h(D(1:t_0))$ is defined similarly to $h(D(1:t))$ as:

$$h(D(1:t_0)) = \left[\max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0), \max_{t_1=t_0}^T \gamma_{t_1}^{\text{min}}(D(1:t_0)) \right\}, \min \left\{ P_{\text{slow}}^{\text{eff-vmax}}(t_0), \min_{t_1=t_0}^T \gamma_{t_1}^{\text{max}}(D(1:t_0)) \right\} \right] \quad (40)$$

$C(P_{\text{slow}}^v(t_0-1))$ is the set of slow-generator output levels that can be reached at time t_0 from its output $P_{\text{slow}}^v(t_0-1)$ at time t_0-1 , which is defined as follows similar to “ $C(\mathbf{P}(t-1))$ ” in (12):

$$C(P_{\text{slow}}^v(t_0-1)) = \{P_{\text{slow}}^v(t_0) : -R_{\text{slow}}^{\text{vdown}}(t_0-1) \leq P_{\text{slow}}^v(t_0) - P_{\text{slow}}^v(t_0-1) \leq R_{\text{slow}}^{\text{vup}}(t_0-1)\}. \quad (41)$$

In order to show that the above algorithm π is robust given $D(1:t)$ and $\mathbf{P}(t)$, we only need to show that $h(D(1:t_0)) \cap C(P_{\text{slow}}^v(t_0-1))$ is nonempty for each $t_0 > t$, and that the value of $P_{\text{fast}}^v(t_0)$ computed by the algorithm π is within the fast generator’s generation limits $[-r_{\text{fast}}^v(t_0), r_{\text{fast}}^v(t_0)]$. We prove the above statement based on the following three claims: As long as the conditions (13)-(14),(17)-(19) hold, we must have

- 1) $h(D(1:t_0)) \neq \emptyset$ for all $t_0 \geq t$.
- 2) $P_{\text{fast}}^v(t_0) = D(t_0) - P_{\text{slow}}^v(t_0) \in [-r_{\text{fast}}^v(t_0), r_{\text{fast}}^v(t_0)]$.
- 3) $h(D(1:t_0)) \cap C(P_{\text{slow}}^v(t_0-1)) \neq \emptyset$ for all $t_0 > t$.

Note that Claim 1 and Claim 2 hold for $t_0 \geq t$. If we let $t_0 = t$, we immediately have $I \neq \emptyset$. In the following three subsections, we will prove the above three claims.

A. Claim 1

In order to prove that $h(D(1:t_0)) \neq \emptyset$, we only need to prove that, when the conditions (13)-(14),(17)-(19) hold, the following four inequalities must be true:

$$1) P_{\text{slow}}^{\text{eff-vmin}}(t_0) \leq P_{\text{slow}}^{\text{eff-vmax}}(t_0); \quad (42)$$

$$2) P_{\text{slow}}^{\text{eff-vmin}}(t_0) \leq \gamma_{t_2}^{\text{max}}(D(1:t_0)), \text{ for all } t_2 \geq t_0; \quad (43)$$

$$3) \gamma_{t_1}^{\text{min}}(D(1:t_0)) \leq P_{\text{slow}}^{\text{eff-vmax}}(t_0), \text{ for all } t_1 \geq t_0; \quad (44)$$

$$4) \gamma_{t_1}^{\text{min}}(D(1:t_0)) \leq \gamma_{t_2}^{\text{max}}(D(1:t_0)), \text{ for all } t_1, t_2 \geq t_0. \quad (45)$$

1) *Proof of Eqn. (42):* We will show that Eqn. (42) is implied by the conditions (13)-(14). According to (15) and

(16), (42) is equivalent to

$$\begin{aligned}
& \max\left\{\max_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\},\right. \\
& \left.\max_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)\right\}\right\} \\
& \leq \min\left\{\min_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vdown}}(s)\right\},\right. \\
& \left.\min_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vup}}(s)\right\}\right\}. \quad (46)
\end{aligned}$$

We first show that

$$\begin{aligned}
& \max_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\} \\
& \leq \min_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vdown}}(s)\right\}.
\end{aligned}$$

Equivalently, we need to show that for any $t_0 < t'_1, t'_2 \leq T$, the following inequality must hold

$$P_{\text{slow}}^{\text{vmin}}(t'_1) - \sum_{s=t_0}^{t'_1-1} R_{\text{slow}}^{\text{vup}}(s) \leq P_{\text{slow}}^{\text{vmax}}(t'_2) + \sum_{s=t_0}^{t'_2-1} R_{\text{slow}}^{\text{vdown}}(s).$$

Note that all the $R_{\text{slow}}^{\text{vup}}(\cdot)$'s and $R_{\text{slow}}^{\text{vdown}}(\cdot)$'s are non-negative. Then, if $t'_1 \leq t'_2$, using Eqn. (13), we have

$$\begin{aligned}
& P_{\text{slow}}^{\text{vmin}}(t'_1) - \sum_{s=t_0}^{t'_1-1} R_{\text{slow}}^{\text{vup}}(s) \\
& \leq P_{\text{slow}}^{\text{vmin}}(t'_1) \leq P_{\text{slow}}^{\text{vmax}}(t'_2) + \sum_{s=t'_1}^{t'_2-1} R_{\text{slow}}^{\text{vdown}}(s) \quad (\text{Eqn. (13)}) \\
& \leq P_{\text{slow}}^{\text{vmax}}(t'_2) + \sum_{s=t_0}^{t'_2-1} R_{\text{slow}}^{\text{vdown}}(s);
\end{aligned}$$

if $t'_1 > t'_2$, using Eqn. (14), we also have

$$\begin{aligned}
& P_{\text{slow}}^{\text{vmin}}(t'_1) - \sum_{s=t_0}^{t'_1-1} R_{\text{slow}}^{\text{vup}}(s) \\
& \leq P_{\text{slow}}^{\text{vmin}}(t'_1) - \sum_{s=t'_2}^{t'_1-1} R_{\text{slow}}^{\text{vup}}(s) \leq P_{\text{slow}}^{\text{vmax}}(t'_2) \quad (\text{Eqn. (14)}) \\
& \leq P_{\text{slow}}^{\text{vmax}}(t'_2) + \sum_{s=t_0}^{t'_2-1} R_{\text{slow}}^{\text{vdown}}(s).
\end{aligned}$$

Using similar arguments, we can also show the following three inequalities:

$$\begin{aligned}
& \max_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\} \\
& \leq \min_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vup}}(s)\right\},
\end{aligned}$$

$$\begin{aligned}
& \max_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)\right\} \\
& \leq \min_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vdown}}(s)\right\},
\end{aligned}$$

and

$$\begin{aligned}
& \max_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)\right\} \\
& \leq \min_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmax}}(t') + \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vup}}(s)\right\}.
\end{aligned}$$

Hence, Eqn. (46) must hold, and thus Eqn. (42) also holds.

2) *Proof of Eqn. (43)*: Before proving Eqn. (43), we first prove the following lemma:

Lemma 12: For any $t \leq t_0 \leq t_1$, we must have

$$P_{\text{slow}}^{\text{eff-vmin}}(t_0) \leq P_{\text{slow}}^{\text{eff-vmin}}(t_1) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s). \quad (47)$$

$$P_{\text{slow}}^{\text{eff-vmin}}(t_0) \geq P_{\text{slow}}^{\text{eff-vmin}}(t_1) - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s). \quad (48)$$

$$P_{\text{slow}}^{\text{eff-vmax}}(t_0) \leq P_{\text{slow}}^{\text{eff-vmax}}(t_1) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s). \quad (49)$$

$$P_{\text{slow}}^{\text{eff-vmax}}(t_0) \geq P_{\text{slow}}^{\text{eff-vmax}}(t_1) - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s). \quad (50)$$

Proof: Here, we only prove (47). Inequalities (48)-(50) can be proved similarly. According to (15), the left-hand-side of (47) is

$$\begin{aligned}
& \max\left\{\max_{t_0 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\},\right. \\
& \left.\max_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)\right\}\right\} \\
& = \max\left\{\max_{t_1 < t' \leq T} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\}, \max_{t_0 < t' \leq t_1} \left\{P_{\text{slow}}^{\text{vmin}}(t')\right.\right. \\
& \left.\left. - \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vup}}(s)\right\}, \max_{t \leq t' \leq t_0} \left\{P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s)\right\}\right\} \quad (51)
\end{aligned}$$

and the right-hand-side of (47) is

$$\begin{aligned}
& \max \left\{ \max_{t_1 < t' \leq T} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_1}^{t'-1} R_{\text{slow}}^{\text{vup}}(s) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s) \right\}, \right. \\
& \quad \left. \max_{t \leq t' \leq t_1} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s) \right\} \right\} \quad (52) \\
& = \max \left\{ \max_{t_1 < t' \leq T} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t_1}^{t'-1} R_{\text{slow}}^{\text{vup}}(s) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vdown}}(s) \right\}, \right. \\
& \quad \max_{t_0 < t' \leq t_1} \left\{ P_{\text{slow}}^{\text{vmin}}(t') + \sum_{s=t_0}^{t'-1} R_{\text{slow}}^{\text{vdown}}(s) \right\}, \\
& \quad \left. \max_{t \leq t' \leq t_0} \left\{ P_{\text{slow}}^{\text{vmin}}(t') - \sum_{s=t'}^{t_0-1} R_{\text{slow}}^{\text{vdown}}(s) \right\} \right\},
\end{aligned}$$

Comparing (51) and (52), we can easily see that (47) holds. \blacksquare

We are now ready to prove Eqn. (43). Using condition (17), Eqn. (47) in Lemma 12, and the fact that $\mathcal{D}_{t_2|D(1:t_0)}$ is a subset of $\mathcal{D}_{t_2|D(1:t)}$, we must have

$$\begin{aligned}
& \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\} \\
& \geq \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t)}} \{D(t_2)\} \\
& \geq P_{\text{slow}}^{\text{eff-vmin}}(t_2) - r_{\text{fast}}^{\text{v-}}(t_2) \\
& \geq P_{\text{slow}}^{\text{eff-vmin}}(t_0) - \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) - r_{\text{fast}}^{\text{v-}}(t_2).
\end{aligned}$$

Thus,

$$\begin{aligned}
& P_{\text{slow}}^{\text{eff-vmin}}(t_0) \\
& \leq \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\} + \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) + r_{\text{fast}}^{\text{v-}}(t_2) \\
& = \gamma_{t_2}^{\text{max}}(D(1:t_0)).
\end{aligned}$$

3) *Proof of Eqn. (44)*: Eqn. (44) can be proved similarly as Eqn. (43). Specifically, using condition (18), Eqn. (50) in Lemma 12, and the fact that $\mathcal{D}_{t_1|D(1:t_0)}$ is a subset of $\mathcal{D}_{t_1|D(1:t)}$, we must have

$$\begin{aligned}
& \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} \\
& \leq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t)}} \{D(t_1)\} \\
& \leq P_{\text{slow}}^{\text{eff-vmax}}(t_1) + r_{\text{fast}}^{\text{v+}}(t_1) \\
& \leq P_{\text{slow}}^{\text{eff-vmax}}(t_0) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) + r_{\text{fast}}^{\text{v+}}(t_1).
\end{aligned}$$

Thus,

$$\begin{aligned}
& P_{\text{slow}}^{\text{eff-vmax}}(t_0) \\
& \geq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) - r_{\text{fast}}^{\text{v+}}(t_1) \\
& = \gamma_{t_1}^{\text{min}}(D(1:t_0)).
\end{aligned}$$

4) *Proof of Eqn. (45)*: According to (19), we must have

$$\begin{aligned}
& r_{\text{fast}}^{\text{v+}}(t_1) + \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) + r_{\text{fast}}^{\text{v-}}(t_2) + \sum_{s=t_0}^{t_2-1} R_{\text{slow}}^{\text{vdown}}(s) \\
& \geq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\}.
\end{aligned}$$

Rearrange the above inequality, we obtain (45).

B. *Claim 2*

Since $P_{\text{slow}}^{\text{v}}(t_0) \in h(D(1:t_0))$, we must have that

$$\gamma_{t_0}^{\text{min}}(D(1:t_0)) \leq P_{\text{slow}}^{\text{v}}(t_0) \leq \gamma_{t_0}^{\text{max}}(D(1:t_0)).$$

Note that $\gamma_{t_0}^{\text{min}}(D(1:t_0)) = D(t_0) - r_{\text{fast}}^{\text{v+}}(t_0)$ and $\gamma_{t_0}^{\text{max}}(D(1:t_0)) = D(t_0) + r_{\text{fast}}^{\text{v-}}(t_0)$, we then have $P_{\text{fast}}^{\text{v}}(t_0) = D(t_0) - P_{\text{slow}}^{\text{v}}(t_0) \in [-r_{\text{fast}}^{\text{v-}}(t_0), r_{\text{fast}}^{\text{v+}}(t_0)]$.

C. *Claim 3*

In order to show that $h(D(1:t_0)) \cap C(P_{\text{slow}}^{\text{v}}(t_0 - 1)) \neq \emptyset$, we only need to show that

$$\begin{aligned}
& \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0), \max_{t_1=t_0}^T \gamma_{t_1}^{\text{min}}(D(1:t_0)) \right\} \\
& \leq P_{\text{slow}}^{\text{v}}(t_0 - 1) + R_{\text{slow}}^{\text{vup}}(t_0 - 1),
\end{aligned}$$

and

$$\begin{aligned}
& \min \left\{ P_{\text{slow}}^{\text{eff-vmax}}(t_0), \min_{t_1=t_0}^T \gamma_{t_1}^{\text{max}}(D(1:t_0)) \right\} \\
& \geq P_{\text{slow}}^{\text{v}}(t_0 - 1) - R_{\text{slow}}^{\text{vdown}}(t_0 - 1).
\end{aligned}$$

Clearly, once we can show that the above two inequalities hold for all $P_{\text{slow}}^{\text{v}}(t_0 - 1) \in h(D(1:t_0 - 1))$, then we must have $h(D(1:t_0)) \cap C(P_{\text{slow}}^{\text{v}}(t_0 - 1)) \neq \emptyset$ for that specific $P_{\text{slow}}^{\text{v}}(t_0 - 1)$. To show the above two inequalities, we only need to prove the following two inequalities:

$$\begin{aligned}
& \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0), \max_{t_1=t_0}^T \gamma_{t_1}^{\text{min}}(D(1:t_0)) \right\} \\
& \leq \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0 - 1), \max_{t_1=t_0-1}^T \gamma_{t_1}^{\text{min}}(D(1:t_0 - 1)) \right\} \\
& \quad + R_{\text{slow}}^{\text{vup}}(t_0 - 1), \quad (53)
\end{aligned}$$

and

$$\begin{aligned}
& \min \left\{ P_{\text{slow}}^{\text{eff-vmax}}(t_0), \min_{t_1=t_0}^T \gamma_{t_1}^{\text{max}}(D(1:t_0)) \right\} \\
& \geq \min \left\{ P_{\text{slow}}^{\text{eff-vmax}}(t_0 - 1), \min_{t_1=t_0-1}^T \gamma_{t_1}^{\text{max}}(D(1:t_0 - 1)) \right\} \\
& \quad - R_{\text{slow}}^{\text{vdown}}(t_0 - 1). \quad (54)
\end{aligned}$$

In the following, we only prove (53). Inequality (54) can be proved similarly. In order to prove (53), we only need to show that the following two inequalities hold:

$$P_{\text{slow}}^{\text{eff-vmin}}(t_0) \leq P_{\text{slow}}^{\text{eff-vmin}}(t_0 - 1) + R_{\text{slow}}^{\text{vup}}(t_0 - 1), \quad (55)$$

$$\gamma_{t_1}^{\text{min}}(D(1:t_0)) \leq \gamma_{t_1}^{\text{min}}(D(1:t_0 - 1)) + R_{\text{slow}}^{\text{vup}}(t_0 - 1), \quad (56)$$

for all $t_1 \geq t_0$.

To see this, note that based on (55) and (56), we must have,

$$\begin{aligned}
& \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0 - 1), \max_{t_1=t_0-1}^T \gamma_{t_1}^{\text{min}}(D(1:t_0 - 1)) \right\} \\
& + R_{\text{slow}}^{\text{vup}}(t_0 - 1) \\
& \geq \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0 - 1), \max_{t_1=t_0}^T \gamma_{t_1}^{\text{min}}(D(1:t_0 - 1)) \right\} \\
& + R_{\text{slow}}^{\text{vup}}(t_0 - 1) \\
& \geq \max \left\{ P_{\text{slow}}^{\text{eff-vmin}}(t_0), \max_{t_1=t_0}^T \gamma_{t_1}^{\text{min}}(D(1:t_0)) \right\},
\end{aligned}$$

which is precisely (53). Finally, to show (55) and (56), note that (55) is a just special case of (48) in Lemma 12. According to (20), (56) can be proved as follows:

$$\begin{aligned}
& \gamma_{t_1}^{\text{min}}(D(1:t_0)) \\
& = \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \sum_{s=t_0}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) - r_{\text{fast}}^{\text{v+}}(t_1) \\
& \leq \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0-1)}} \{D(t_1)\} - \sum_{s=t_0-1}^{t_1-1} R_{\text{slow}}^{\text{vup}}(s) - r_{\text{fast}}^{\text{v+}}(t_1) \\
& + R_{\text{slow}}^{\text{vup}}(t_0 - 1) \\
& = \gamma_{t_1}^{\text{min}}(D(1:t_0 - 1)) + R_{\text{slow}}^{\text{vup}}(t_0 - 1).
\end{aligned}$$

This completes the proof.

APPENDIX E

DEMAND SPLITTING IDEA AND ITS SUBOPTIMALITY

In this section, we study a naive implementation of the demand splitting idea. In this naive demand splitting (NDS) approach, we divide the ‘‘uncertain part’’ (see Section IV-B2 for the detailed definition) of the net-demand into N_g parts, each of which corresponds to a certain fraction of the total uncertainty and is then assigned to a generator. The main difference between the NDS approach and the VDS approach is that the NDS approach does not pair each generator with a virtual fast generator. In fact, the NDS approach can be viewed as a special case of the VDS approach, where $R_{\text{fast,g}}^{\text{v+}}(t) = R_{\text{fast,g}}^{\text{v-}}(t) = r_{\text{fast,g}}^{\text{v+}}(t) = r_{\text{fast,g}}^{\text{v-}}(t) = 0$. Next, we will show that this NDS approach can have very poor performance.

We construct an example based on the two-generator case in Section IV-A. Let $P_{\text{slow}}^{\text{vmin}}(t) = 0$, $P_{\text{slow}}^{\text{vmax}}(t) = P_M$, $R_{\text{slow}}^{\text{vdown}}(t) = R_{\text{slow}}^{\text{vup}}(t) = R$, $r_{\text{fast}}^{\text{v-}}(t) = 0$, $r_{\text{fast}}^{\text{v+}}(t) = r$ be the generator parameters, where P_M, R, r are constants. As for the net-demand, we assume that $0 \leq D(t) \leq D_M = P_M + r$ and $|D(t) - D(s)| \leq a|t - s| + \Delta$, where D_M, a, Δ are constants. The parameters D_M, a, Δ determines the uncertainty set \mathcal{D} .

We first determine the conditions for $\mathcal{D} \in \Lambda$ in the above scenario. In this case, the constraints (13)-(18) always hold, and the constraint (19) can be rewritten as follows:

$$\begin{aligned}
& (t_1 - t_0)R + (t_2 - t_0)R + r \\
& \geq \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]}} \left\{ \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D(t_1)\} - \right. \\
& \quad \left. \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D(t_2)\} \right\} \quad (57) \\
& = \min\{D_M, a(t_1 - t_0 + t_2 - t_0) + 2\Delta\}, \text{ for all } t_1, t_2 \geq t_0.
\end{aligned}$$

In the following, we will only focus on one case, where $a = R, r = 2\Delta = 2(n - 1)R, D_M = P_M + r = nr$. It is easy to verify that this parameter setting satisfies the constraint (57). Thus, this two-generator system is robust for the uncertainty set \mathcal{D} . We then apply the NDS approach in this parameter setting, and show that the performance can be very poor.

Specifically, let η_s be the fraction of the uncertain part of the demand that is assigned to the slow generator. Then, based on Eqn. (31) in Section IV-B2, the total demand allocated to the slow generator is

$$D_{\text{slow}}(t_0) = P_{\text{slow}}^{\text{main}}(t_0) + \eta_s(D(t) - D^{\text{main}}(t_0)).$$

We need to make sure that this slow generator alone can serve all possible realizations of the demand $D_{\text{slow}}(t_0)$. Consider the constraint (19) in Theorem 11. Letting $t = 0, t_0 = t_1 = t_2 - 1$, we obtain

$$\begin{aligned}
R & \geq \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]}} \left\{ \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D_{\text{slow}}(t_1)\} - \right. \\
& \quad \left. \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D_{\text{slow}}(t_2)\} \right\} \\
& = P_{\text{slow}}^{\text{main}}(t_0) - P_{\text{slow}}^{\text{main}}(t_0 + 1) - \eta_s(D^{\text{main}}(t_0) \\
& \quad - D^{\text{main}}(t_0 + 1)) + \eta_s \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]}} \left\{ D(t_0) - \right. \\
& \quad \left. \min_{D(t_0+1) \in \mathcal{D}_{t_0+1|D(1:t_0)}} \{D(t_0 + 1)\} \right\} \\
& = P_{\text{slow}}^{\text{main}}(t_0) - P_{\text{slow}}^{\text{main}}(t_0 + 1) - \eta_s(D^{\text{main}}(t_0) \\
& \quad - D^{\text{main}}(t_0 + 1)) + \eta_s(a + \Delta). \quad (58)
\end{aligned}$$

In the last step of (58), we have used the parameter setting that $|D(t) - D(s)| \leq a|t - s| + \Delta$ for any t and s . Similarly, letting $t = 0, t_0 = t_1 - 1 = t_2$, we obtain

$$\begin{aligned}
R & \geq \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]}} \left\{ \max_{D(t_1) \in \mathcal{D}_{t_1|D(1:t_0)}} \{D_{\text{slow}}(t_1)\} - \right. \\
& \quad \left. \min_{D(t_2) \in \mathcal{D}_{t_2|D(1:t_0)}} \{D_{\text{slow}}(t_2)\} \right\} \\
& = P_{\text{slow}}^{\text{main}}(t_0 + 1) - P_{\text{slow}}^{\text{main}}(t_0) - \eta_s(D^{\text{main}}(t_0 + 1) \\
& \quad - D^{\text{main}}(t_0)) + \eta_s \max_{D(1:t_0) \in \mathcal{D}_{[1,t_0]}} \left\{ -D(t_0) + \right. \\
& \quad \left. \max_{D(t_0+1) \in \mathcal{D}_{t_0+1|D(1:t_0)}} \{D(t_0 + 1)\} \right\} \\
& = P_{\text{slow}}^{\text{main}}(t_0 + 1) - P_{\text{slow}}^{\text{main}}(t_0) - \eta_s(D^{\text{main}}(t_0 + 1) \\
& \quad - D^{\text{main}}(t_0)) + \eta_s(a + \Delta). \quad (59)
\end{aligned}$$

Sum up the two inequalities (58) and (59), we obtain

$$\eta_s \leq \frac{R}{a + \Delta} = \frac{1}{n}. \quad (60)$$

Similarly, let η_f denote the fraction of the uncertain part of the demand that is assigned to the fast generator. Then, based on Eqn. (31) in Section IV-B2, the total demand allocated to the fast generator is

$$D_{\text{fast}}(t_0) = P_{\text{fast}}^{\text{main}}(t_0) + \eta_f(D(t_0) - D^{\text{main}}(t_0)).$$

We need to make sure that this fast generator alone can serve all possible realizations of the demand $D_{\text{fast}}(t_0)$. Consider the

constraints (17) and (18) in Theorem 11. Letting $t = 0$, we obtain

$$\begin{aligned} 0 - 0 &\leq \min_{D(t_0) \in \mathcal{D}_{t_0}} \{D_{\text{fast}}(t_0)\} \\ &= P_{\text{fast}}^{\text{main}}(t_0) + \eta_f(0 - D^{\text{main}}(t_0)), \end{aligned} \quad (61)$$

and

$$\begin{aligned} 0 + r &\geq \max_{D(t_0) \in \mathcal{D}_{t_0}} \{D_{\text{fast}}(t_0)\} \\ &= P_{\text{fast}}^{\text{main}}(t_0) + \eta_f(D_M - D^{\text{main}}(t_0)). \end{aligned} \quad (62)$$

Subtract the inequality (62) by the inequality (61), we obtain $r \geq \eta_f D_M$, and thus,

$$\eta_f \leq \frac{r}{D_M} = \frac{1}{n}. \quad (63)$$

In order for this two-generator system to serve all the uncertain part of the demand $D(t)$, we need $\eta_s + \eta_f = 1$. However, based on (60) and (63), $\eta_s + \eta_f$ can be much smaller than 1 if n is large. Recall that this two-generator system is in fact robust according to (57). This example thus indicates that, by restricting splitting the demand onto different generators, the performance can be quite poor.

We highlight the insight in the above example. The slow generator has large capacity. However, if its ramping rate is low, we cannot send a large η_s to the slow generator when the demand changing rate is large. On the contrary, the fast generator has large ramping rate. However, because of its low capacity, we cannot send a large η_f either. Intuitively, if we pair a slow generator with a fast generator, they complement each other's constraints, and thus can be robust for a larger uncertainty set. This observation thus motivates us to consider demand splitting to paired set of slow+fast generators. Specifically, in the next section, we will propose an improved version of the demand splitting approach, in which a fast generator and a slow generator are grouped into a VGP (virtual generator pair) in order to accommodate a larger fraction of uncertainty.

APPENDIX F ROBUSTNESS OF THE VDS-ED ALGORITHM

We study the robustness of the VDS-ED algorithm in this section. Our objective is to prove the following theorem:

Theorem 13: Given an uncertainty set \mathcal{D} , suppose that for any t and demand history $D(1:t)$, $\mathcal{F}^{\text{VDS}}(D(1:t)) \neq \emptyset$. Then, the VDS-ED algorithm is robust for the uncertainty set \mathcal{D} , i.e., for any t and any demand history $D(1:t)$, the dispatch decision $\mathbf{P}(t)$ chosen by the VDS-ED algorithm must satisfy the constraints (4)-(7).

Since $\mathbf{P}(t)$ is chosen within the set $C(\mathbf{P}(t-1))$, the constraint (5) must be satisfied. Further, the following lemma states that $\mathcal{F}^{\text{VDS}}(D(1:t)) \subseteq A_t(D(t))$.

Lemma 14: For any t and any demand history $D(1:t)$, we have $\mathcal{F}^{\text{VDS}}(D(1:t)) \subseteq A_t(D(t))$.

Proof: See Appendix F-A. \blacksquare

Based on the above discussion, if we can choose a $\mathbf{P}(t)$ from the set $\mathcal{F}^{\text{VDS}}(D(1:t)) \cap C(\mathbf{P}(t-1))$, then $\mathbf{P}(t)$ must

satisfy the constraints (4)-(7). Hence, in order to show that the VDS-ED algorithm is robust for the uncertainty set \mathcal{D} , we only need to show that $\mathcal{F}^{\text{VDS}}(D(1)) \neq \emptyset$ and $\mathcal{F}^{\text{VDS}}(D(1:t)) \cap C(\mathbf{P}(t-1)) \neq \emptyset$ for all $t > 1$.

$\mathcal{F}^{\text{VDS}}(D(1)) \neq \emptyset$ holds trivially according to the assumption in Theorem 13. Next, we focus on the proof of $\mathcal{F}^{\text{VDS}}(D(1:t+1)) \cap C(\mathbf{P}(t)) \neq \emptyset$ for all $t \geq 1$.

We have assumed that $\mathcal{F}^{\text{VDS}}(D(1:t)) \neq \emptyset$. Consider an arbitrary $\mathbf{P}(t) \in \mathcal{F}^{\text{VDS}}(D(1:t))$. Based on the definition (36) of $\mathcal{F}^{\text{VDS}}(D(1:t))$, there must exist $Z(t) = \{\eta_{b,g}, R_{\text{fast},g}^{\text{v-}}(t_0), R_{\text{fast},g}^{\text{v+}}(t_0), r_{\text{fast},g}^{\text{v-}}(t_0), r_{\text{fast},g}^{\text{v+}}(t_0), P_{\text{VGP},g}^{\text{main}}(t_0), t_0 \geq t, b \in \mathcal{B}, g \in \mathcal{G}\}$ satisfying all admissible constraints in Sec. IV-B1 and IV-B2 including (25)-(30) and (32)-(33), and there exists $(P_{\text{slow},g}^{\text{v}}(t), P_{\text{fast},g}^{\text{v}}(t)) \in \mathcal{F}_g^{Z(t)}(D(1:t)) \neq \emptyset, g \in \mathcal{G}$, satisfying (34) and (35).

Note that each VGP g can serve all possible realizations of the net-demand $D_g(t+1:T)$. Then, given the demand $D_g(t+1)$ (computed from $D(t+1)$ using (31)), each VGP can choose a dispatch decision

$$(P_{\text{slow},g}^{\text{v}}(t+1), P_{\text{fast},g}^{\text{v}}(t+1)) \in \mathcal{F}_g^{Z(t)}(D(1:t+1)), \quad (64)$$

such that $-R_{\text{slow},g}^{\text{vdown}}(t) \leq P_{\text{slow},g}^{\text{v}}(t+1) - P_{\text{slow},g}^{\text{v}}(t) \leq R_{\text{slow},g}^{\text{vup}}(t)$. Let $\mathbf{P}(t+1)$ be a physical generator dispatch decision, such that $P_g(t+1)$ and $(P_{\text{slow},g}^{\text{v}}(t+1), P_{\text{fast},g}^{\text{v}}(t+1))$ satisfy (34) and (35). According to (36), we know $\mathbf{P}(t+1) \in \mathcal{F}^{\text{VDS}}(D(1:t+1))$. Further,

$$\begin{aligned} &P_g(t+1) - P_g(t) \\ &= P_g(t+1) - P_{\text{slow},g}^{\text{v}}(t+1) + P_{\text{slow},g}^{\text{v}}(t+1) \\ &\quad - P_{\text{slow},g}^{\text{v}}(t) + P_{\text{slow},g}^{\text{v}}(t) - P_g(t) \\ &\leq R_{\text{fast},g}^{\text{v+}}(t+1) + R_{\text{slow},g}^{\text{vup}}(t) + R_{\text{fast},g}^{\text{v-}}(t) \\ &= R_g. \end{aligned}$$

Similarly, we can also prove that $P_g(t+1) - P_g(t) \geq -R_g$. Therefore, $\mathbf{P}(t+1) \in C(\mathbf{P}(t))$. As a result $\mathcal{F}^{\text{VDS}}(D(1:t+1)) \cap C(\mathbf{P}(t))$ contains at least $\mathbf{P}(t+1)$, and thus is not empty.

Since $\mathcal{F}^{\text{VDS}}(D(1:t+1)) \cap C(\mathbf{P}(t)) \neq \emptyset$, it is always possible to pick the most economic dispatch decision in $\mathcal{F}^{\text{VDS}}(D(1:t+1)) \cap C(\mathbf{P}(t))$ in the VDS-ED algorithm. Further, since $\mathcal{F}^{\text{VDS}}(D(1:t+1)) \subseteq A_{t+1}(D(t+1))$ according to Lemma 14, the most economic dispatch decision chosen in the VDS-ED algorithm must satisfy the constraints (4)-(7). This completes the proof of Theorem 13.

A. Proof of Lemma 14

Proof: Consider an arbitrary $\mathbf{P}(t) \in \mathcal{F}^{\text{VDS}}(D(1:t))$. Then, based on the definition (36) of $\mathcal{F}^{\text{VDS}}(D(1:t))$, there must exist $Z(t) = \{\eta_{b,g}, R_{\text{fast},g}^{\text{v-}}(t_0), R_{\text{fast},g}^{\text{v+}}(t_0), r_{\text{fast},g}^{\text{v-}}(t_0), r_{\text{fast},g}^{\text{v+}}(t_0), P_{\text{VGP},g}^{\text{main}}(t_0), t_0 \geq t, b \in \mathcal{B}, g \in \mathcal{G}\}$ satisfying all admissible constraints in Sec. IV-B1 and IV-B2 including (25)-(30) and (32)-(33), and there exists $(P_{\text{slow},g}^{\text{v}}(t), P_{\text{fast},g}^{\text{v}}(t)) \in \mathcal{F}_g^{Z(t)}(D(1:t)) \neq \emptyset, g \in \mathcal{G}$, satisfying (34) and (35).

First, let $t_0 = t$ in (32) and (33) (note that $\mathcal{D}_{t|D(1:t)}$ only contains a single demand value $D(t)$), we obtain

$$\left| \sum_{b=1}^{N_b} S_{l,b} \left(D_b(t) - \sum_{g \in \mathcal{G}_b} D_g(t) \right) \right| \leq \text{TL}_l.$$

Noting that $\sum_{g \in \mathcal{G}_b} D_g(t) = \sum_{g \in \mathcal{G}_b} (P_{\text{slow},g}^v(t) + P_{\text{fast},g}^v(t)) = \sum_{g \in \mathcal{G}_b} P_g(t)$ according to (35), we then obtain that $\mathbf{P}(t)$ satisfies constraint (7).

Second, sum up the equation (35) over all $b \in \mathcal{B}$, we obtain

$$\begin{aligned} \sum_{g \in \mathcal{G}} P_g(t) &= \sum_{b \in \mathcal{B}} \sum_{g \in \mathcal{G}_b} P_g(t) = \sum_{b \in \mathcal{B}} \sum_{g \in \mathcal{G}_b} D_g(t) \\ &= \sum_{g \in \mathcal{G}} \left(P_{\text{VGP},g}^{\text{main}}(t) + \sum_{b \in \mathcal{B}} \eta_{b,g} (D_b(t) - D_b^{\text{main}}(t)) \right) \\ &= \sum_{b \in \mathcal{B}} D_b^{\text{main}}(t) + \sum_{b \in \mathcal{B}} (D_b(t) - D_b^{\text{main}}(t)) \sum_{g \in \mathcal{G}} \eta_{b,g} \\ &= \sum_{b \in \mathcal{B}} D_b(t). \end{aligned}$$

Therefore, $\mathbf{P}(t)$ satisfies the constraint (6).

Finally, since $(P_{\text{slow},g}^v(t), P_{\text{fast},g}^v(t)) \in \mathcal{F}_g^{Z(t)}(D(1:t))$, we must have

$$P_{\text{slow},g}^{\text{vmin}}(t) = P_{\text{slow},g}^{\text{eff-vmin}}(t) \leq P_{\text{slow},g}^v(t) \leq P_{\text{slow},g}^{\text{eff-vmax}}(t) \leq P_{\text{slow},g}^{\text{vmax}}(t).$$

Then, according to (34), we have

$$P_g(t) \leq P_{\text{slow},g}^v(t) + R_{\text{fast},g}^{v+}(t) \leq P_{\text{slow},g}^{\text{vmax}}(t) + R_{\text{fast},g}^{v+}(t) = P_g^{\text{maximum}},$$

and

$$P_g(t) \geq P_{\text{slow},g}^v(t) - R_{\text{fast},g}^{v-}(t) \geq P_{\text{slow},g}^{\text{vmin}}(t) - R_{\text{fast},g}^{v-}(t) = P_g^{\text{minimum}}.$$

Hence, $\mathbf{P}(t)$ also satisfies the constraint (4).

Recall the definition (9) of $A_t(D(t))$, we then have $\mathbf{P}(t) \in A_t(D(t))$. Since $\mathbf{P}(t)$ is chosen arbitrarily from $\mathcal{F}^{\text{VDS}}(D(1:t))$, we must have $\mathcal{F}^{\text{VDS}}(D(1:t)) \subseteq A_t(D(t))$. ■

APPENDIX G ADDITIONAL SIMULATION DATA

In Section V, we have demonstrated that the standard ED algorithm will fail when the scaling factor equals 1.9, but the VDS-ED algorithm can still ensure system safety. In this section, we present more detailed simulation data, so that readers can easily see the dynamics in the system, and how the VDS-ED algorithm improves the dispatch decisions of the ED algorithm.

In Table IV, we record all the generators' outputs before the ED algorithm fails. (Note that the "*" next to a output level at 7:45am indicates that the output level reaches the lower limit of the corresponding generator.) The ED algorithm fails at 8:00am, when net demand (load minus renewable) drops from 5130MW (7:45am) to 5007MW (8:00am). In order to balance the demand at time 8am, those generators that operate higher than their minimum need to ramp down. Here, based on the schedule of the ED algorithm at 7:45am, only 3 of the Type-B generators (those at levels 422MW, 568MW and 600MW) can ramp down. Recall that the ramping rate of the Type-B generators is 4MW/min. Hence, each Type-B generator can ramp down by at most 60MW in 15 minutes, if it has not hit its lower generation limit. Given the dispatch decision at time 7:45am (Table IV), it is easy to check that the demand cannot be met at time 8am. Specifically, the generator with output level 422MW can only ramp down by 2MW, and thus

the total output can ramp down by at most $2 + 60 + 60 = 122\text{MW}$, which is smaller than total demand drop $5130 - 5007 = 123\text{MW}$. In summary, even though there may exist another dispatch decision at 7:45am that could have been safe for 8:00am (see the dispatch decision at 7:45am of the VDS algorithm in Table V. There are 3 type-B generators whose levels are 515MW, 474MW, and 600MW, and thus can ramp down by at most $60 + 54 + 60 = 174 > 123\text{MW}$), the ED algorithm may choose an unsafe dispatch decision, because it fails to account for the future uncertainty.

Readers may notice that the energy generation costs of the ED algorithm and the VDS-ED algorithm are the same at most time-slots. This indicates that in these time-slots, there happen to be an economic dispatch decision that is also robust for the future uncertainty. Due to this reason, the VDS-ED algorithm often achieves low costs comparable to the ED algorithm. However, there will be cases when the cost of the VDS-ED algorithm is higher than that of the ED algorithm at some time-slots, e.g., 6:45 am. This occurs when none of the most economic decision is safe for the future. Specifically, from Fig. 2(b) and Fig. 2(c), we can see that the net-demand drops quite rapidly around 6:45am. In this case, the VDS-ED algorithm will pick a slightly less economic dispatch decision, in order to ensure the system safety in the future.

TABLE IV. DISPATCH DECISIONS OF THE STANDARD ED ALGORITHM (SCALING FACTOR: 1.9)

Generators	6:00 am	6:15 am	6:30 am	6:45 am	7:00 am	7:15 am	7:30 am	7:45 am
Type-A #1 (bus 1)	600MW	750MW	715MW	606MW	600MW	600MW	600MW	600MW(*)
Type-A #2 (bus 1)	600MW	600MW	600MW	600MW	600MW	600MW	600MW	600MW(*)
Type-B #1 (bus 1)	600MW	600MW	600MW	600MW	540MW	480MW	420MW	420MW(*)
Type-B #2 (bus 1)	600MW	600MW	600MW	600MW	540MW	480MW	482MW	422MW
Type-B #3 (bus 1)	600MW	600MW	600MW	600MW	540MW	529MW	469MW	420MW(*)
Type-B #1 (bus 4)	600MW	600MW	600MW	600MW	569MW	509MW	569MW	568MW
Type-B #2 (bus 4)	600MW	600MW	600MW	600MW	540MW	600MW	600MW	600MW
Type-C #1 (bus 4)	490MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)
Type-C #2 (bus 4)	300MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)
Type-C #3 (bus 4)	300MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)
Type-C #4 (bus 4)	300MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)
Type-C #5 (bus 4)	300MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)
13 Type-D generators	0	0	0	0	0	0	0	0
Total Demand	5890MW	5850MW	5815MW	5706MW	5429MW	5298MW	5240MW	5130MW
Cost	18339\$	18187\$	18058\$	17646\$	16948\$	16621\$	16475\$	16199\$

TABLE V. DISPATCH DECISIONS OF THE VDS-ED ALGORITHM (SCALING FACTOR: 1.9)

Generators	6:00 am	6:15 am	6:30 am	6:45 am	7:00 am	7:15 am	7:30 am	7:45 am	8:00 am	8:15 am	8:30 am	8:45 am
Type-A #1 (bus 1)	600MW	600MW	600MW	600MW	600MW	600MW	600MW	600MW(*)	600MW	600MW	600MW	600MW
Type-A #2 (bus 1)	600MW	600MW	715MW	600MW	600MW	600MW	600MW	600MW(*)	600MW	600MW	600MW	600MW
Type-B #1 (bus 1)	600MW	600MW	600MW	540MW	480MW	480MW	420MW	420MW(*)	420MW	420MW	420MW	420MW
Type-B #2 (bus 1)	600MW	600MW	600MW	600MW	600MW	575MW	575MW	515MW	455MW	420MW	420MW	480MW
Type-B #3 (bus 1)	600MW	600MW	600MW	600MW	540MW	480MW	465MW	474MW	420MW	476MW	467MW	420MW
Type-B #1 (bus 4)	600MW	600MW	600MW	600MW	583MW	583MW	600MW	600MW	540MW	540MW	480MW	480MW
Type-B #2 (bus 4)	600MW	600MW	600MW	586MW	526MW	480MW	480MW	420MW(*)	472MW	428MW	480MW	421MW
Type-C #1 (bus 4)	300MW	450MW	300MW	300MW	300MW	300MW	300MW	300MW(*)	300MW	300MW	300MW	300MW
Type-C #2 (bus 4)	300MW	300MW	400MW	380MW	300MW	300MW	300MW	300MW(*)	300MW	300MW	300MW	300MW
Type-C #3 (bus 4)	300MW	300MW	300MW	300MW	300MW	300MW	300MW	300MW(*)	300MW	300MW	300MW	300MW
Type-C #4 (bus 4)	300MW	300MW	300MW	380MW	300MW	300MW	300MW	300MW(*)	300MW	300MW	300MW	300MW
Type-C #5 (bus 4)	490MW	369MW	315MW	300MW	300MW	300MW	300MW	300MW(*)	300MW	300MW	300MW	300MW
13 Type-D generators	0	0	0	0	0	0	0	0	0	0	0	0
Total Demand	5890MW	5850MW	5815MW	5706MW	5429MW	5298MW	5240MW	5130MW	5007MW	4984MW	4967MW	4921MW
Cost	18339\$	18187\$	18058\$	17739\$	16948\$	16621\$	16475\$	16199\$	15893\$	15835\$	15792\$	15677\$