DTU Library

An Efficient Robust Solution to the Two-Stage Stochastic Unit Commitment Problem

Blanco, Ignacio; Morales González, Juan Miguel

An Efficient Robust Solution to the Two-Stage Stochastic Unit Commitment Problem

Ignacio Blanco and Juan M. Morales, *Senior Member, IEEE,*

Abstract—This paper provides a reformulation of the scenario-based two-stage unit commitment problem under uncertainty that allows finding unit-commitment plans that perform reasonably well both in expectation and for the worst case. The proposed reformulation is based on partitioning the sample space of the uncertain factors by clustering the scenarios that approximate their probability distributions. The degree of conservatism of the resulting unit-commitment plan (that is, how close it is to the one provided by a purely robust or stochastic unit-commitment formulation) is controlled by the number of partitions into which the said sample space is split. To efficiently solve the proposed reformulation of the unit-commitment problem under uncertainty, we develop two alternative parallelization and decomposition schemes that rely on a column-and-constraint generation procedure. Finally, we analyze the quality of the solutions provided by this reformulation for a case study based on the IEEE 14-node power system and test the effectiveness of the proposed parallelization and decomposition solution approaches on the larger IEEE 3-Area RTS-96 power system.

Keywords—*Stochastic and robust unit commitment, column-and-constraint generation, parallel computing, clustering, scenario reduction.*

Nomenclature

The notation used throughout the paper is stated below for quick reference. Other symbols are defined as required.

A. Indexes and Sets

T	Set of time periods t.
N	Set of nodes n.
G	Set of conventional generation units g.
F	Set of stochastic power production units f.
L	Set of loads l.
Ω	Set of scenarios ω, ranging from 1 to λ.
P	Set of partitions p, ranging from 1 to k.
Ω_p	Set of scenarios ω in partition p.
F_n	Set of stochastic power production units located at node n.
L_n	Set of loads connected at node n.
G_n	Set of conventional generation units located at node n.
M_n	Set of nodes $m \in N$ that are connected to node n by a transmission line.
Ω'_p	Reduced set of scenarios ω in partition p.

B. Parameters

C_g^F, C_g^V	Fixed/variable production cost of conventional generation unit g.
C_g^{SU}, C_g^{SD}	Start-up/Shut-down cost of conventional generation unit g.
$L_{l,t}$	Demand for load l at time t.
RU_g, RD_g	Ramp-up/Ramp-down rate for conventional generation unit g.
UT_g, DT_g	Minimum-up/Minimum-down time for unit g
L_g^{UP}, L_g^{DW}	Number of time periods conventional generation unit g must be online/offline counting from $t = 1$.
IS_g	Initial status of unit g, equal to 1 if online at $t = 0$ and 0, otherwise.
ON_g, OFF_g	Number of time periods unit g has been online/offline prior to $t = 1$.
$X_{n,m}$	Reactance of line $n - m$.
$F_{n,m}^{max}$	Maximum flow capacity of line $n - m$.
P_g^{max}, P_g^{min}	Maximum/minimum power production of conventional generation unit g
P_g^{SU}, P_g^{SD}	Maximum starting-up/shutting-down power production of conventional generation unit g.
P_g^{IS}	Power output of conventional unit g at $t = 0$.
C^L	Cost of involuntary load curtailment.
$W_{f,t,\omega}$	Power production from stochastic generation unit f at time t in scenario ω.
π_ω	Probability of scenario ω.
ρ_p	Weight associated with partition p.

C. First-stage variables

$u_{g,t}$	Binary variable equal to 1 if unit g is online at time t and 0, otherwise.
$y_{g,t}/z_{g,t}$	Binary variable equal to 1 if unit g is starting up/shutting down at time t and 0, otherwise.

D. Second-stage variables

$P_{g,t,\omega}$	Power produced by conventional generation unit g in scenario ω at time t.
$L_{l,t,\omega}^{SH}$	Power curtailment from load l in scenario ω at time t.
$W_{f,t,\omega}^{SP}$	Power curtailment from stochastic power production unit f in scenario ω at time t.
$\delta_{n,t,\omega}$	Voltage angle at node n, time t and scenario ω.
α	Auxiliary variable used in the scenario-based robust unit commitment formulation
θ_p	Auxiliary variable used in the hybrid unit commitment formulation

I. Introduction

The increasing reliance on partly unpredictable renewable power supply has prompted the revision of the procedures used for power system operations. This is the case, for example, of the tool used by system operators to decide the commitment of power plants, that is, to solve the so-called *unit commitment problem* (UC). Two-stage stochastic

programming [1] and robust optimization [2] have become the most popular and explored techniques of optimization under uncertainty to improve unit-commitment decisions in terms of both cost-efficiency and system reliability.

The formulation and solution of the unit commitment problem using either stochastic programming or robust optimization—the result of which is typically referred to as *stochastic* and *robust unit commitment*, respectively—has been subject of numerous studies by the scientific community; see, for instance, [3]–[11], among many others and variants.

Essentially, the stochastic unit commitment problem (SUC) makes use of a probabilistic model for the uncertain input factors such as demand, equipment failures and partly-predictable renewable power production to minimize a certain quantile of the induced system cost distribution or its expectation. Most often than not, this probabilistic model is approximated by a set of scenarios that describe plausible realizations of such random factors. In order for the stochastic solution to be reliable, the amount of scenarios that need to be considered must be large, which may render an intractable optimization problem, or carefully generated, which motivates the topic of *scenario reduction techniques* [12]–[14]. Furthermore, the probabilistic model from which these scenarios may be drawn may carry, in itself, some level of uncertainty as well.

In contrast, the robust unit commitment problem (RUC) seeks a commitment plan that allows the system to withstand the worst-case realization of the uncertain factors at a minimum cost. While this approach saves the decision-maker from having to probabilistically characterize these factors, it may yield too conservative solutions as the worst-case scenario rarely occurs.

In recent years, several methods have been proposed to make decisions under uncertainty that perform *relatively well* under the premises of *both* the stochastic and the robust approaches, that is, in expectation and for the worst case. Illustrative examples of these methods can be found in [15]–[18], where hybrid stochastic-robust solution strategies are developed for optimal air-quality and municipal solid-waste management, electricity trading for power microgrids and energy contracting for a portfolio of renewable power generation technologies, respectively. What makes all these solution strategies *hybrid* is that some of the uncertain parameters are assumed to follow certain probability distributions, while others are solely known to belong to some uncertainty sets.

Within the context of the unit commitment problem, we highlight the work in [19] and [20]. More specifically, the authors in [19] propose a mathematical formulation that delivers the unit-commitment plan that minimizes a user-controlled weighted sum of the expected and the worst-case costs. The solution approach introduced in [20], even if presented as a method to tackle the stochastic unit commitment problem, seeks to determine a unit-commitment plan that is robust against an *ambiguous* probability distribution of renewable energy generation, an ambiguity that is the result of the always limited availability of data and that is modeled in practice as a vector of imperfectly known probabilities. Thus, as the amount of historical data increases, the ambiguity of such a probability distribution diminishes and so does the need for robustness and the degree of conservatism of the stochastic unit-commitment solution. This approach can also be regarded as a form of *distributionally robust optimization* [21].

Our work shares with [19] and [20] the aim of finding a solution to the stochastic unit commitment problem that is robust in some sense, but our motivation and the methodology we propose to this end are essentially different. We assume that the probability distributions of the uncertain parameters—in our case, the wind power production—are known, but that, as it normally occurs in practice, computational tractability only allows us to solve the stochastic unit commitment problem for a scenario-based approximation of such distributions. In principle, we shall consider a large number of scenarios for this approximation to be accurate enough. In any case, we group these scenarios using a clustering technique—for instance, the *k-means clustering algorithm* [24], which has been reported to feature good performance in similar contexts [25], [26]—. Each of the so-obtained clusters is referred to as a *partition*. We then formulate and solve a two-stage unit-commitment problem that minimizes the expected value of the system operating costs, where the expectation is taken over the collection of worst-case scenarios within each partition. The probability assigned to each of these worst-case scenarios is equal to the probability of the partition it belongs to, which is, in turn, computed by summing up the probabilities of the scenarios that form part of the partition in question.

For convenience, we employ the term *hybrid unit commitment problem* and the acronym HUC to refer to the proposed reformulation of the UC problem. This reformulation brings two major advantages, namely:

1) It allows finding solutions to the two-stage unit commitment problem with a decreasing degree of conservatism by increasing the number of partitions. In fact, if only one partition is considered, the HUC delivers the (scenario-based) robust unit-commitment solution (whereby "scenario-based", we mean the solution given by the robust formulation of the two-stage unit-commitment problem in the particular case that the uncertainty is modeled as a finite set of atoms, outcomes or scenarios). In contrast, if the number of partitions is made to coincide with the number of scenarios, the HUC solution boils down to the stochastic unit-commitment plan. Furthermore, as we show later, the computational time required to find the HUC solution increases with the number of partitions considered, with the solution algorithm being very fast for a small number of them. This provides a practical way to adjust the level of conservatism of the unit-commitment solution depending on time availability.

2) It is amenable to decomposition and paralellization in various ways and levels and, hence, it can be efficiently solved. Indeed, based on the column-and-constraint generation procedure described in [9], we provide and compare two alternative decomposition schemes to solve the HUC problem. These two schemes basically differ in whether the worst-case scenarios within each partition are identified independently for each partition or not.

The remainder of this paper is organized as follows. Section II begins by providing mathematical formulations for the scenario-based two-stage stochastic and robust unit commitment problems, in that order, and finishes with the formulation of the proposed hybrid unit commitment problem. Furthermore, in this section we explain how we use a clustering method to construct the partitions in the HUC model and how these can be employed to control the degree of conservatism of the resulting unit-commitment plan. Section III introduces the proposed parallelization and decomposition strategies to solve the HUC problem. Section

IV analyzes and discusses results from two case studies based on standard IEEE power systems. Finally, in Section V the main conclusions of our work are summarized, including possible avenues for future research.

II. MATHEMATICAL FORMULATION

In the two-stage unit commitment problem under uncertainty, decision variables are divided into two groups. The first group constitutes the commitment plan itself and consists of the 0/1 variables $u_{g,t}$, $y_{g,t}$, $z_{g,t}$, which determine the on/off status, the start-up, and the shutdown of generating unit g in time period t, respectively. These decisions are to be made, in general, one day in advance of the actual delivery of electricity and, in any case, before the realization of the uncertain factors. In this paper, we consider for simplicity that the system uncertainty stems only from the wind power production, which is modeled as a finite set Ω of scenarios $W_{f,t,\omega}$ with $\omega \in \Omega$.

The second-stage decision variables, namely, $P_{g,t,\omega}$, $L_{l,t,\omega}^{SH}$, $W_{f,t,\omega}^{SP}$ and $\delta_{n,t,\omega}$ determine the economic dispatch of the conventional generating units, the amount of load that is involuntarily shed, the amount of wind power production that is curtailed, and the voltage angles at the network nodes, respectively. These variables adapt to the specific realization of the uncertainty and as such, are augmented with the scenario index ω.

We start by providing the mathematical formulation of the two-stage *stochastic* unit commitment problem. In all cases, we consider that the marginal production cost of the wind generation is zero.

A. Two-Stage Stochastic Unit Commitment (SUC)

The two-stage stochastic unit commitment problem can be formulated as follows:

$$\underset{\mathcal{H},\mathcal{W}}{\text{minimize}} \quad \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t} + C_g^{SU} y_{g,t} + C_g^{SD} z_{g,t} \right) \quad (1)$$

$$+ \sum_{\omega \in \Omega} \pi_\omega \left[\sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega} + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH} \right]$$

$$\text{s.t.} \quad y_{g,t} - z_{g,t} = u_{g,t} - u_{g,t-1} \quad (2)$$
$$(\forall g, \forall t \in \{2,...,T\})$$

$$y_{g,t} - z_{g,t} = u_{g,t} - IS_g \quad (3)$$
$$(\forall g, \forall t \in \{1\})$$

$$y_{g,t} + z_{g,t} \leq 1 \quad (4)$$
$$(\forall g, \forall t \in \{1,...,T\})$$

$$u_{g,t} = IS_g \quad (5)$$
$$(L_g^{UP} + L_g^{DW} > 0, \forall g, \forall t \leq L_g^{UP} + L_g^{DW})$$

$$\sum_{\tau=t-UT_g+1}^{t} y_{g,\tau} \leq u_{g,t} \quad (6)$$
$$(\forall g, \forall t > L_g^{UP} + L_g^{DW})$$

$$\sum_{\tau=t-DT_g+1}^{t} z_{g,\tau} \leq 1 - u_{g,t} \quad (7)$$
$$(\forall g, \forall t > L_g^{UP} + L_g^{DW})$$

$$\sum_{g \in G_n} P_{g,t,\omega} - \sum_{l \in L_n} L_{l,t} + \sum_{l \in L_n} L_{l,t,\omega}^{SH} + \sum_{f \in F_n} W_{f,t,\omega}$$
$$- \sum_{f \in F_n} W_{f,t,\omega}^{SP} = \sum_{m \in M_n} \frac{(\delta_{n,t,\omega} - \delta_{m,t,\omega})}{X_{n,m}} \quad (8)$$
$$(\forall n, \forall t, \forall \omega \in \Omega)$$

$$P_{g,t,\omega} \leq P_g^{max} u_{g,t} \quad (9)$$
$$(\forall g, \forall t, \forall \omega \in \Omega)$$

$$P_{g,t,\omega} \geq P_g^{min} u_{g,t} \quad (10)$$
$$(\forall g, \forall t, \forall \omega \in \Omega)$$

$$P_{g,t,\omega} \leq (P_g^{IS} + RU_g) u_{g,t} \quad (11)$$
$$(\forall g, \forall t \in \{1\}, \forall \omega \in \Omega)$$

$$P_{g,t,\omega} \geq (P_g^{IS} - RD_g) u_{g,t} \quad (12)$$
$$(\forall g, \forall t \in \{1\}, \forall \omega \in \Omega)$$

$$P_{g,t,\omega} - P_{g,t-1,\omega} \leq (2 - u_{g,t-1} - u_{g,t}) P_g^{SU} \quad (13)$$
$$+ (1 + u_{g,t-1} - u_{g,t}) RU_g$$
$$(\forall g, \forall t \in \{2,...,T\}, \forall \omega \in \Omega)$$

$$P_{g,t-1,\omega} - P_{g,t,\omega} \leq (2 - u_{g,t-1} - u_{g,t}) P_g^{SD} \quad (14)$$
$$+ (1 - u_{g,t-1} + u_{g,t}) RD_g$$
$$(\forall g, \forall t \in \{2,...,T\}, \forall \omega \in \Omega)$$

$$L_{l,t,\omega}^{SH} \leq L_{l,t} \quad (15)$$
$$(\forall l, \forall t, \forall \omega \in \Omega)$$

$$W_{f,t,\omega}^{SP} \leq W_{f,t,\omega} \quad (16)$$
$$(\forall f, \forall t, \forall \omega \in \Omega)$$

$$- F_{n,m}^{max} \leq \frac{(\delta_{n,t,\omega} - \delta_{m,t,\omega})}{X_{n,m}} \leq F_{n,m}^{max} \quad (17)$$
$$(\forall n, m \in M_n, \forall t, \forall \omega \in \Omega)$$

$$P_{g,t,\omega}, L_{l,t,\omega}^{SH}, W_{f,t,\omega}^{SP} \geq 0 \quad (18)$$
$$(\forall g, \forall l, \forall f, \forall t, \forall \omega \in \Omega)$$

$$u_{g,t}, y_{g,t}, z_{g,t} \in \{0,1\} \quad (19)$$
$$(\forall g, \forall t)$$

where $\mathcal{H} = \{u_{g,t}, y_{g,t}, z_{g,t}\}$ and $\mathcal{W} = \left\{ P_{g,t,\omega}, L_{l,t,\omega}^{SH}, W_{f,t,\omega}^{SP}, \delta_{n,t,\omega} : \omega \in \Omega \right\}$ are the sets of here-and-now and wait-and-see decisions, respectively. Furthermore, following [28], the initial state conditions are given by

$$IS_g = \begin{cases} 1 & \text{if } ON_g > 0 \\ 0 & \text{if } ON_g = 0 \end{cases}$$

$$L_g^{UP} = \min\{T, (UP_g - ON_g)IS_g\}$$

$$L_g^{DW} = \min\{T, (DT_g - OFF_g)(1 - IS_g)\}$$

Problem (1)–(19) takes the form of a standard two-stage unit commitment formulation, which is similar, to a large extent, to those provided in the numerous works on the topic, see, for instance, [6], [27] and references therein. The objective is to minimize the expected system operating

cost (1), which is made up of the no-load, start-up, shutdown, and variable production costs of the conventional generating units, and the cost of involuntarily load curtailment, in that order. Equations (2)–(4) model the changes in the on/off-commitment status of the power plants as these are started up or shutdown throughout the scheduling horizon, while (5)–(7) impose their minimum up- and down-time requirements. Equalities (8) constitute the set of nodal power balance equations according to a DC power flow model. The maximum and minimum power outputs of the generating units are enforced by (9) and (10), respectively, and their ramping limits through (11)–(14), as in [8] and [10]. The sets of inequalities (15) and (16) limit the involuntary load curtailment and the wind power spillage to the eventual power that is consumed and the eventual wind power that is produced, respectively. The set of equations (17) guarantee compliance with the transmission capacity limits. Finally, constraints (18) and (19) constitute variable declarations.

B. Two-Stage Robust Unit Commitment (RUC)

The two-stage robust unit commitment problem can be written as follows:

$$\min_{\mathcal{H}} \ \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t} + C_g^{SU} y_{g,t} + C_g^{SD} z_{g,t} \right) + \mathcal{Q}(\mathcal{H}) \tag{20}$$

$$\text{s.t. } (2) - (7), (19) \tag{21}$$

where

$$\mathcal{Q}(\mathcal{H}) = \max_{\omega \in \Omega} \min_{\mathcal{W}} \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega} + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH} \tag{22}$$

$$\text{s.t. } (8) - (18) \tag{23}$$

In the particular case that the uncertainty set Ω is comprised of a *finite* number of atoms, outcomes or scenarios ω, the two-stage robust unit-commitment problem (20)–(23) can be equivalently recast as follows:

$$\underset{\mathcal{H},\mathcal{W},\alpha}{\text{minimize}} \ \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t} + C_g^{SU} y_{g,t} + C_g^{SD} z_{g,t} \right) + \alpha \tag{24}$$

$$\text{s.t. } \alpha \geq \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega} + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH}, \ \ \forall \omega \in \Omega \tag{25}$$

$$(2) - (19) \tag{26}$$

For convenience and ease of reference, the set of equations (24)–(26) is hereafter referred to as *the scenario-based formulation of the two-stage robust unit-commitment problem*. Note that the auxiliary variable α equals the worst-case dispatch cost at the optimum. This variable is bounded from below by a finite set of linear constraints (25), one per scenario, that involve the second-stage decision variables $P_{g,t,\omega}$ and $L_{l,t,\omega}^{SH}$. Thus, the objective of the problem (24)–(26) is to minimize the total system operating cost for the worst-case scenario of the uncertainty.

In the following section, we introduce the proposed hybrid formulation of the two-stage unit commitment problem under uncertainty.

C. Hybrid Unit Commitment Problem (HUC)

Let us now split the set Ω into k partitions with $P = \{1, ..., k\}$ being the partition set. We then define the series

of subsets $\Omega_1, \ldots, \Omega_p, \ldots, \Omega_k$, with $\Omega_i \bigcap \Omega_j = \emptyset$ for all $i \neq j$ and $\Omega_1 \bigcup \ldots \Omega_p \bigcup \ldots \Omega_k = \Omega$, such that Ω_p is linked to partition $p \in P$. Furthermore, each partition $p \in P$ is assigned a probability $\rho_p \geq 0$ such that $\sum_{p \in P} \rho_p = 1$.

The proposed hybrid two-stage unit commitment problem writes as follows:

$$\min_{\mathcal{H}} \ \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t} + C_g^{SU} y_{g,t} + C_g^{SD} z_{g,t} \right) + \mathcal{R}(\mathcal{H}) \tag{27}$$

$$\text{s.t. } (2) - (7), (19) \tag{28}$$

where

$$\mathcal{R}(\mathcal{H}) = \sum_{p \in P} \rho_p \left(\max_{\omega \in \Omega_p} \min_{\mathcal{W}} \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega} \right.$$
$$\left. + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH} \right) \tag{29}$$

$$\text{s.t.} \qquad (8) - (18) \tag{30}$$

In the particular case that the uncertainty set Ω is *finite*, that is, $\Omega = \{1, ..., \lambda\}$, where λ is the total number of possible outcomes or scenarios, problem (27)–(30) can be equivalently reformulated as follows:

$$\underset{\mathcal{H},\mathcal{W},\theta_p}{\text{minimize}} \ \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t} + C_g^{SU} y_{g,t} + C_g^{SD} z_{g,t} \right) +$$
$$+ \sum_{p \in P} \rho_p \theta_p \tag{31}$$

$$\text{s.t. } \theta_p \geq \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega} + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH} \tag{32}$$
$$(\forall p \in P, \ \forall \omega \in \Omega_p)$$

$$(2) - (19) \tag{33}$$

where Ω_p is comprised of all the scenarios $\omega \in \Omega$ that belong to partition $p \in P$ and where the probability ρ_p assigned to partition p is computed as the sum of the probabilities of the scenarios that form part of it, that is,

$$\rho_p = \sum_{\omega \in \Omega_p} \pi_\omega \quad \forall p \in P \tag{34}$$

The objective (31) is then to minimize the expected system operating cost over the scenarios that deliver the worst-case dispatch cost within each partition.

Indeed, the auxiliary variable θ_p, one per partition $p \in P$, equals the worst-case dispatch cost within partition p, in a similar way as the auxiliary variable α does in the robust unit commitment formulation (24)–(26) for the whole set of scenarios Ω. This way, problem (31)–(33) is expected to yield a unit-commitment plan that is "in between" the robust and the stochastic unit-commitment solutions in terms of the expected and the worst-case system operating cost. Furthermore, the closeness of the HUC solution to the stochastic and robust unit-commitment plans, and consequently its degree of conservatism, are controlled by the number k of partitions or clusters into which the scenarios are grouped. Indeed, if the number of partitions equals the number of scenarios, that is, $k = \lambda$, the HUC model (31)–(33) reduces to (1)–(19) and the stochastic solution is obtained. In contrast, if only one single partition is considered ($k = 1$), we have that $\Omega_1 = \Omega$ and therefore, problem (31)–(33) boils down to the scenario-based robust unit-commitment formulation (24)–(26). As a

result, the HUC solution coincides with the robust solution in such a case.

Hence, we can increase the degree of conservatism of the HUC solution by diminishing the number of partitions, and vice versa. For $1 < k < \lambda$, however, how efficiently and quickly the HUC solution transits from the robust to the stochastic unit-commitment plan, as k increases, depends on the performance of the clustering technique. Later on in this paper, we evaluate and compare the performance of the proposed HUC formulation when solved using a range of different clustering methods. More specifically, we consider three non-hierarchical clustering techniques, namely, the k-means, k-medoids, and k-shape algorithms, which are used to group a data set into a given number k of clusters. In brief, these algorithms assign each scenario $\omega \in \Omega$ to the partition Ω_p, $p \in P$, with the nearest mean (k-means), the closest representative scenario (k-medoids) [22], or the most similar scenario-shape using a cross-correlation measure (k-shape) [23]. In particular, the k-means algorithm [24] has been reported to showcase a good performance in other related applications, e.g., for solving a probabilistic production cost model in [25] and a transmission and generation expansion planning problem in [26]. Furthermore, we also test an agglomerative hierarchical clustering method [22], where the dissimilarity of two clusters is measured as the maximum of the pairwise distances of the scenarios in the clusters. This method produces a hierarchy of partitions whereby the two nearest clusters merge into a new one as one moves up the hierarchy. Therefore, the top level in the hierarchy consists of one single cluster that comprises the complete set of scenarios.

III. SOLUTION STRATEGY: PARALLELIZATION AND DECOMPOSITION

It is well known that the unit commitment problem is mixed-integer, NP-hard, and generally requires long solution times. This is especially true for realistic instances of the two-stage unit commitment problem under uncertainty. In the following we describe two alternative parallelization-and-decomposition schemes that we have designed to efficiently solve the proposed HUC formulation (31)–(33). For ease of exposition, we divide this description in two parts. In the first one, we explain how problem (31)–(33) can be decomposed per partition and scenario, while in the second part we elaborate on how the solution to the decomposed problem can be parallelized.

A. Problem Decomposition via Column-and-constraint Generation

We provide and discuss below two alternative ways to decompose the HUC problem (31)–(33), which we present as two variants of the same *Scenario Partition and Decomposition Algorithm* (SPDA).

1) Scenario Partition and Decomposition Algorithm— Variant 1 (SPDA1): Let us consider a certain partition $p \in P$ that comprises the subset of scenarios Ω_p. Note that, for determining the optimal solution to the HUC problem (31)– (33), we only need those (hopefully few) scenarios in Ω_p that deliver the worst-case dispatch cost over partition p for *any* feasible unit-commitment plan. Identifying those scenarios, however, may be computationally very costly. Instead, we describe below a procedure to find a subset $\Omega'_p \subset \Omega_p$ under which the HUC formulation provides a unit commitment solution close to the one given under the full scenario set

Ω_p. Furthermore, the proposed algorithm builds the reduced sets Ω'_p, $p \in P$, in parallel for each partition p.

To build Ω'_p from Ω_p, the latter being the outcome of a certain clustering algorithm, we develop a master-subproblem decomposition scheme *per partition* based on the column-and-constraint generation procedure described in [9]. In the sequel we will refer to this decomposition scheme as *Primal Cut Algorithm* after the solution strategy introduced in [10] whereby the master problem is gradually enlarged with the addition of cuts expressed in terms of the primal variables.

Each master problem (one per partition) is a mixed-integer programming problem that involves both first-stage and second-stage decision variables and that has the following form at iteration i of the column-and-constraint generation algorithm:

$$\underset{\mathcal{H}^i, \mathcal{W}^i, \theta_p}{\text{minimize}} \quad \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i \right) + \theta_p \tag{35}$$

$$\text{s.t.} \quad (2) - (7), (19) \tag{36}$$

$$\theta_p \geq \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega}^i + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH,i}, \quad \forall \omega \in \Omega_p'^i \tag{37}$$

$$(8) - (18), \quad \forall \omega \in \Omega_p'^i \tag{38}$$

where $\mathcal{H}^i = \left\{ u_{g,t}^i, y_{g,t}^i, z_{g,t}^i \right\}$ and $\mathcal{W}^i = \left\{ P_{g,t,\omega}^i, L_{l,t,\omega}^{SH,i}, W_{f,t,\omega}^{SP,i}, \delta_{n,t,\omega}^i : \omega \in \Omega_p'^i \right\}$. Note that $\Omega_p'^0 = \emptyset$. As the algorithm proceeds, $\Omega_p'^i$ is augmented with those possibly few scenarios $\omega \in \Omega_p$ that are needed to reconstruct the partition-worst-case recourse cost as a function of the first-stage decision variables $u_{g,t}^i$, $y_{g,t}^i$, and $z_{g,t}^i$ in the form of (37)–(38).

Constraint (37) can be interpreted as a primal cut, as compared to those cuts that are constructed from dual information, as it is the case, for example, of a standard Benders cut.

The subproblems are linear programming problems (LP) that determine the second-stage decision variables $P_{g,t,\omega}^i$, $L_{l,t,\omega}^{SH,i}$, $W_{f,t,\omega}^{SP,i}$, and $\delta_{n,t,\omega}^i$ with $u_{g,t}^i$, $y_{g,t}^i$, and $z_{g,t}^i$ fixed at the values given by the master problem. A subproblem like (39)–(40) is solved for each scenario $\omega \in \Omega_p$.

$$\underset{\mathcal{W}_\omega^i}{\text{minimize}} \quad \sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega}^i + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega}^{SH,i} \tag{39}$$

$$\text{s.t.} \quad (8) - (18) \tag{40}$$

where $\mathcal{W}_\omega^i = \left\{ P_{g,t,\omega}^i, L_{l,t,\omega}^{SH,i}, W_{f,t,\omega}^{SP,i}, \delta_{n,t,\omega}^i \right\}$.

The scenario ω' for which the associated subproblem (39)–(40) yields the highest dispatch cost or is infeasible is used to construct a set of primal constraints in the form of (37)–(38) that is added to the master problem by setting $\Omega_p'^{i+1} = \Omega_p'^i \bigcup \{\omega'\}$. It is worth noticing, however, that subproblem infeasibility is not a concern in our case due to the possibility of shedding load and spilling wind.

One instance of the primal cut algorithm is run for each partition $p \in P$ in parallel. Each of these instances works, therefore, with one master problem and a number of subproblems equal to the number of scenarios in each partition, that is, equal to $card(\Omega_p)$. Furthermore, each instance of the algorithm concludes by delivering the set

of selected scenarios $\Omega'_p \subset \Omega_p$ for partition p. The last step of our solution strategy consists then in solving the HUC problem (31)–(33) where Ω_p is replaced with the reduced scenario set Ω'_p.

We describe below how this solution strategy proceeds step by step.

1) Choose the number k of partitions and apply a clustering method to the complete set of scenarios Ω in order to assign each scenario to a certain partition p.

2) Create one instance of the primal cut algorithm for each partition $p \in P$.

3) Initialization: Set $i = 0$ and $\Omega_p'^0 = \emptyset$.

4) Solve the master problem (MP). Return the optimal solution found by the branch-and-cut algorithm and denote this solution by $(u_{g,t}^i, y_{g,t}^i, z_{g,t}^i)$. Calculate a lower bound LB as $\sum_{t\in T}\sum_{g\in G}(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i) + \theta_p$.

5) Solve the subproblems (SP) with the first-stage decision variables fixed at $(u_{g,t}^i, y_{g,t}^i, z_{g,t}^i)$. Once the SP are solved, the scenario ω' associated with the subproblem that yields the highest dispatch cost is identified and included into the reduced set $\Omega_p'^{i+1}$, i.e., $\Omega_p'^{i+1} = \Omega_p'^i \bigcup \{\omega'\}$. Compute an upper bound UB as $\sum_{t\in T}\sum_{g\in G}\left(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i\right) + \sum_{t\in T}\sum_{g\in G} C_g^V P_{g,t,\omega'}^i + \sum_{t\in T}\sum_{l\in L} C^L L_{l,t,\omega'}^{SH,i}$.

6) Convergence check: If $\mid UB - LB \mid \leq \epsilon$, being ϵ a user-specified tolerance value, the iterative process stops. If $\mid UB - LB \mid > \epsilon$, then set $i := i+1$ and go to step 4.

7) Once all the instances of the primal cut algorithm have converged, the HUC problem (31)–(33) is solved for all $p \in P$ and for all $\omega \in \Omega'_p$. The reduced set Ω'_p is made up of those scenarios $\omega \in \Omega_p$ that determine the worst-case dispatch cost within partition p.

A pseudocode for the proposed decomposition scheme is provided in Algorithm 1. For ease of notation, let x (x^i) denote the vector of first-stage variables (at iteration i).

Algorithm 1 Scenario Partition and Decomposition Algorithm: Variant 1 (**SPDA1**)

1: Choose k and apply k-means to Ω.
2: **for** all $p \in P$ **do**
3:      Set $i := 0$ and $\Omega_p'^0 = \emptyset$
4:      **repeat**
5:          **Solve** Master Problem
6:          Return optimal solution x^i
7:          Compute Lower Bound LB
8:          **Set** $x := x^i$ and **solve** SP $\forall \omega \in \Omega_p$
9:          Compute Upper Bound UB
10:         Identify worst-case scenario ω'
11:         **Set** $\Omega_p'^{i+1} := \Omega_p'^i \bigcup \{\omega'\}$
12:         **Set** $i := i+1$
13:      **until** $\mid UB - LB \mid \leq \epsilon$
14:      **Set** $\Omega'_p := \Omega_p'^{i-1}$
15: **end for**
16: **Solve** HUC replacing Ω_p with Ω'_p, $\forall p$

Notice that SPDA1 works in a similar way to a scenario reduction technique that retains the most detrimental scenarios in terms of system operating cost. This confers robustness to the solution of the proposed HUC problem. Moreover, the last command line in SPDA, which involves solving the HUC model for the reduced scenario sets Ω'_p, $\forall p \in P$,

could be carried out as well via further decomposition (see, for instance, [29]), although this possibility has not been explored in this paper.

2) Scenario Partition and Decomposition Algorithm—Variant 2 (SPDA2): Algorithm SPDA1 is inspired from the idea that only a few scenarios in the full set Ω_p may be needed to determine the worst-case dispatch cost for partition p and for *any* feasible unit-commitment plan. Based on this, Algorithm SPDA1 generates one instance of the Primal Cut Algorithm *per partition*.

The reader should notice, however, that, in order to compute the minimum expected system operating cost over the scenarios that deliver the worst-case dispatch cost within each partition, one might not need to calculate the partition-worst-case dispatch cost for *any* feasible unit-commitment plan and therefore, one might not need all those scenarios that algorithm SPDA1 aims to identify, but possibly a smaller number of them. With this in mind, we construct a second variant of our Scenario Partition and Decomposition Algorithm, which we denote SPDA2.

Unlike its first variant, SPDA2 only generates one single instance of the Primal Cut Algorithm, that is, a single master-subproblem scheme. The master problem is a mixed-integer programming problem that takes the following form at iteration i of the column-and-constraint generation procedure:

$$\underset{\mathcal{H}^i, \mathcal{W}^i, \theta_p}{\text{minimize}} \quad \sum_{t\in T}\sum_{g\in G}\left(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i\right) + $$
$$+ \sum_{p\in P} \rho_p \theta_p \tag{41}$$

$$\text{s.t.} \quad (2)-(7), \ (19) \tag{42}$$

$$\theta_p \geq \sum_{t\in T}\sum_{g\in G} C_g^V P_{g,t,\omega}^i + \sum_{t\in T}\sum_{l\in L} C^L L_{l,t,\omega}^{SH,i}, \tag{43}$$
$$(\forall p \in P, \ \forall \omega \in \Omega_p'^i)$$

$$(8)-(18), \quad \forall p \in P, \ \forall \omega \in \Omega_p'^i \tag{44}$$

where $\mathcal{H}^i = \{u_{g,t}^i, y_{g,t}^i, z_{g,t}^i\}$ and $\mathcal{W}^i = \{P_{g,t,\omega}^i, L_{l,t,\omega}^{SH,i}, W_{f,t,\omega}^{SP,i}, \delta_{n,t,\omega}^i \ : \ \omega \in \Omega_p'^i, \ p \in P\}$, with $\Omega_p'^0 = \emptyset$ for all $p \in P$.

The subproblems are linear programming problems analogous to (39)–(40). At every iteration i of the algorithm, the master problem (41)–(44) produces a tentative unit-commitment plan $\mathcal{H}^i$ that is fed into the subproblems (39)–(40). The scenarios $\{\omega'_1, \omega'_2, \ldots, \omega'_p, \ldots, \omega'_k\}$, one per partition, that result in the highest dispatch cost within each partition are used to generate primal cuts in the form of (43)–(44) that are inserted into the master problem by setting $\Omega_p'^{i+1} = \Omega_p'^i \bigcup \{\omega'_p\}$, $\forall p \in P$.

We provide next a step-by-step description of SPDA2.

1) Choose the number k of partitions and apply a clustering method to the full set of scenarios Ω in order to assign each scenario to a certain partition p.

2) Create one instance of the primal cut algorithm.

3) Initialization: Set $i = 0$ and $\Omega_p'^0 = \emptyset, \forall p \in P$.

4) Solve the master problem (MP). Return the optimal solution found by the branch-and-cut algorithm and denote this solution by $(u_{g,t}^i, y_{g,t}^i, z_{g,t}^i)$. Calculate a lower bound LB as $\sum_{t\in T}\sum_{g\in G}(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i) + \sum_{p\in P} \rho_p \theta_p$.

5) Solve the subproblems (SP) with the first-stage decision variables fixed at $(u_{g,t}^i, y_{g,t}^i, z_{g,t}^i)$. Once the SP are solved, the scenarios $\{\omega_1', \omega_2', \ldots, \omega_p', \ldots, \omega_k'\}$ associated with the subproblems that yield the highest dispatch cost in each partition p are identified and included into the reduced sets $\Omega_p'^{i+1}$, $p \in P$, i.e., $\Omega_p'^{i+1} = \Omega_p'^i \bigcup \{\omega_p'\}, \forall p \in P$. Compute an upper bound UB as

$$UB = \sum_{t \in T} \sum_{g \in G} \left(C_g^F u_{g,t}^i + C_g^{SU} y_{g,t}^i + C_g^{SD} z_{g,t}^i \right)$$
$$+ \sum_{p \in P} \rho_p \left(\sum_{t \in T} \sum_{g \in G} C_g^V P_{g,t,\omega_p'}^i + \sum_{t \in T} \sum_{l \in L} C^L L_{l,t,\omega_p'}^{SH,i} \right).$$

6) Convergence check: If $| UB - LB | \leq \epsilon$, being ϵ a user-specified tolerance value, the iterative process stops and $(u_{g,t}^i, y_{g,t}^i, z_{g,t}^i)$ is returned as the optimal solution to the HUC problem. If $| UB - LB | > \epsilon$, then set $i := i + 1$ and go to step 4.

A pseudocode of this solution strategy is provided in Algorithm 2 below.

Algorithm 2 Scenario Partition and Decomposition Algorithm: Variant 2 (**SPDA2**)

1: Choose k and apply k-means to Ω.
2: Set $i := 0$ and $\Omega_p'^0 = \emptyset$ for all $p \in P$
3: **repeat**
4:     **Solve** Master Problem
5:     Return optimal solution x^i
6:     Compute Lower Bound LB
7:     **Set** $x := x^i$ and **solve** SP $\forall \omega \in \Omega$
8:     Compute Upper Bound UB
9:     Identify worst-case scenario ω_p' in each $p \in P$
10:    **Set** $\Omega_p'^{i+1} := \Omega_p'^i \bigcup \{\omega_p'\}$ for all $p \in P$
11:    **Set** $i := i + 1$
12: **until** $| UB - LB | \leq \epsilon$
13: **Return** $x := x^{i-1}$ as the solution to HUC

The two variants of the proposed Scenario Partition and Decomposition Algorithm have several key differences, namely:

1) SPDA2 guarantees convergence to the HUC solution that is optimal under the partition P of the full scenario set Ω, while SPDA1 does not. Indeed, SPDA1 is heuristic, because it does not ensure that the scenarios the algorithm selects from each partition p are enough to deliver the worst-case dispatch cost for that partition under *any* first-stage solution. Our numerical experiments show, however, that SPDA1 performs very well in the sense that it provides unit commitment solutions that are optimal for the full HUC problem (31)–(33). Furthermore, it should be noticed that SPDA1 can be used to warm-start SPDA2.

2) Let Ω_1' and Ω_2' be the reduced set of scenarios retained by the master problems of SPDA1 and the sole master problem of SPDA2 at the end of the algorithms, respectively. In all our simulations (see the case study in Section IV) it holds that $card(\Omega_2') \leq card(\Omega_1') \leq card(\Omega) = \lambda$.

3) At every iteration, SPDA2 solves one single master problem, while SPDA1 solves k, but smaller master problems, with $k = card(P)$.

4) SPDA1 must solve the HUC problem for the reduced scenario set Ω_1' at the end of the algorithm.

Differences 2-4 above determine how these two variants compare in computational terms. In particular, it is expected that, since SPDA1 works with smaller master problems at each iteration, but must solve a possibly larger HUC problem in the end, SPDA1 will perform better than SPDA2 when the number of partitions is kept sufficiently low. This is corroborated by our numerical experiments.

B. Parallelization of the Solution Algorithm

In the case of SPDA1 both the outer "for-loop" and the solution to the ω-indexed subproblems in the pseudocode of "Algorithm 1" are amenable to parallelization. For this purpose, we make use of the DTU High Performance Computing (HPC) Facility [30]. We create k jobs, each representing an instance of the primal cut algorithm for each of the k partitions into which we divide the scenario set Ω. These jobs are simultaneously submitted to the HPC Cluster, where they are concurrently executed, as there is no need for communication in between the workers (nodes or cores).

We submit each of the k jobs to a different node, using the same amount of resources per node. Within every node, the subproblems are solved in a multi-threaded environment using the Gather-Update-Solve-Scatter Facility in GAMS [31]. This tool allows treating each subproblem, one per scenario in the partition under consideration, as a different parametrization of the same linear programming model, which is then generated only once by GAMS. Likewise, the solutions to all the subproblems (or portions thereof) are retrieved back to GAMS in a single transaction.

In the case of SPDA2, the solution to the ω-indexed subproblems are also solved in the same way as in SPDA1, that is, by means of the the Gather-Update-Solve-Scatter Facility in GAMS [31].

IV. CASE STUDIES

In the following the unit-commitment solution provided by the proposed HUC formulation is tested on the IEEE 14-node power system [32] and the IEEE 3-Area RTS-96 system. The latter power system is also used to evaluate and compare the performance of the SPDA algorithms in Sections IV-C and IV-D.

The IEEE 14-node system comprises 14 nodes, 5 generators, 20 lines and 11 loads. We also add one wind farm to node 5, whose power production is modeled by the ten scenarios provided in Table IV of the Appendix.

The IEEE 3-Area RTS-96 system consists of 72 nodes, 96 generators, 107 lines, and 51 loads. Besides, we add 15 wind farms of 200-MW capacity each and location given by Table V in the Appendix. Thus, the wind power capacity represents 29% of the total generating capacity installed in the power system. The technical characteristics of the generating units, the demand and the transmission lines are available online [33]. The wind power production scenarios used for this case study come from [34], where the spatio-temporal dependencies of wind power generation are considered. More specifically, the study in [34] provides 100 scenarios of wind power production that were generated for 15 control areas and 43 lead times in western Denmark. However, for this work, only 50 equiprobable scenarios and 24 lead times are considered.

We set the MIP tolerance gap to 0 in all the simulations pertaining to the IEEE 14-node system, while we allow for

TABLE I: Results of the comparison between Zhao model and the proposed HUC formulation.

UCP	CCD [$]	# Partitions	ZWF	ETC [$]	WCTC [$]
1	62569	6 - 10	0.8 - 1	286602	311534
2	63505	1 - 5	0 - 0.8	287131	307030

a MIP tolerance gap of up to $3 \cdot 10^{-4}$ (0.03%) in those numerical experiments carried out on the IEEE 3-Area RTS-96 system.

A. Evaluation of the Effect of the Number of Partitions

We consider the IEEE 14-node system and compare the unit-commitment plans resulting from the proposed HUC formulation with those obtained using the method proposed in [19], which we refer to as *Zhao model* hereafter. This method minimizes a convex combination of the expected and the worst-case costs, where the convex combination is defined by the decision-maker through a weighting factor α. This way, Zhao model allows for a Pareto-efficient control of the degree of conservatism of the resulting unit-commitment solution in terms of the expected and worst-case system operating costs. For this reason, we use Zhao model here as an ideal benchmark against which we compare the ability of our proposed HUC formulation to find unit-commitment plans with a varying degree of conservatism by changing the number of partitions. The comparison is conducted for a number of partitions in the HUC problem ranging from 1 to 10 and a number of values for the weighting factor in Zhao model varying from 0 to 1.

The outcome of this comparison is that both Zhao model and our HUC formulation deliver the same only two unit-commitment plans, which we denote by UCP1 and UCP2. We measure the quality of these two different unit-commitment solutions in terms of both the expected and the worst-case system operating cost. The results are collated in Table I. The columns of this table provide the unit-commitment plan (UCP), the commitment cost (CCD), the number of partitions considered in our HUC formulation, the value of the weighting factor used in Zhao model (ZWF), the expected total cost (ETC) and the worst-case total cost (WCTC), in that order. Costs are given in US dollars. Furthermore, recall that the pure stochastic and robust unit-commitment solutions are obtained for a number of partitions equal to 10 and 1 in the proposed HUC formulation, and for values of the weighting factor in Zhao model equal to 1 and 0, respectively.

For this specific test we have used the k-means clustering method to define the partitions.

We can see from Table I that UCP2 is more conservative than UCP1 and that there exists a one-to-one correspondence between the unit-commitment solutions provided by Zhao model and those given by the proposed HUC formulation. Thus, the proposed HUC formulation offers an alternative way to control the degree of conservatism of the resulting unit-commitment solution by way of the number of partitions. Besides, the *efficient frontier* (made up of Pareto-efficient points) is, in general, non-linear and non-continuous. Consequently, tuning the solution conservatism through the number of partitions, which is a natural number that ranges from 1 to the number of scenarios considered, is more practical than doing it through a weighting factor α that is a real number belonging to the closed real interval $[0, 1]$.

B. Evaluation of the Impact of the Clustering Technique

Next we analyze how the proposed HUC formulation performs when different clustering techniques are used to construct the partitions. For this purpose, we conduct some numerical tests on the IEEE 3-Area RTS-96 system.
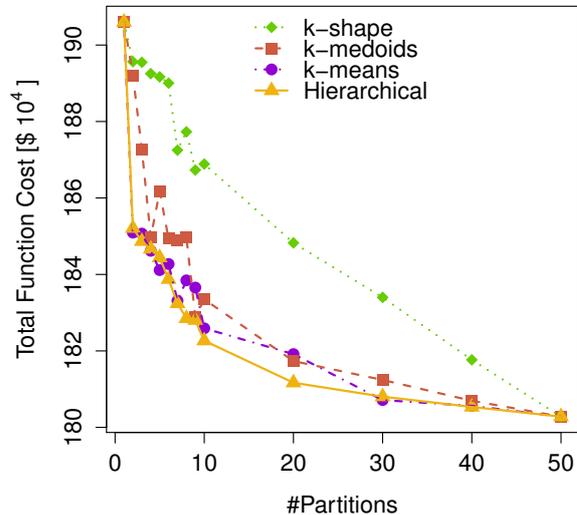

Fig. 1: HUC total cost as a function of the number of partitions and for four different clustering techniques.

Fig. 1 represents the objective function value (31) of the HUC formulation as the number of partitions increases. Each plot corresponds to a different clustering technique. As mentioned above, we consider the non-hierarchical clustering algorithms k-means, k-shape, and k-medoids, and an agglomerative hierarchical method.

In general terms, all the four clustering techniques prompt HUC solutions for which the HUC objective function value decreases as the number of partitions increases, as expected, but they do so at different speeds. In particular, the k-means and the hierarchical procedures achieve large reductions in the HUC cost with very few partitions. In practice, this means that these two clustering algorithms are able to prompt solutions from the HUC formulation that are close to the pure stochastic one with few partitions. This is clearly an advantage for two reasons at least: first, because our objective is, all in all, to approach the stochastic unit-commitment solution in a robust manner, that is, while keeping the worst-case cost under control; second, because, as we shall see later, the time required to solve the proposed HUC formulation decreases as so does the number of partitions.

In contrast, the plot corresponding to the k-shape method transits slowly from the pure robust solution to the pure stochastic one, which means that this methods requires a higher number of partitions to reduce the degree of conservatism of the UC solution.

It is interesting to note that, unlike the hierarchical clustering technique, the k-means, k-shape and k-medoids algorithms do not guarantee a monotone decrease in the HUC cost with the number of partitions. This is due to the non-hierarchical nature of these clustering methods, whereby the partitions they produce for $k-1$, k and $k+1$ clusters bear no relation, that is, they do not necessarily result from splitting or merging existing clusters.

The results that follow have been obtained using the hierarchical clustering technique, although very similar results are also obtained if the k-means algorithm is used.

C. Evaluation of the Decomposition Schemes

In this section we use the case study based on the IEEE 3-Area RTS-96 system to investigate and compare three different approaches for solving the proposed HUC formulation, namely, the two variants of the proposed Scenario Partition and Decomposition Algorithm (SPDA1 and SPDA2) and a solution strategy that merely consists in directly solving the "raw" HUC problem (without decomposition). All these alternative solution approaches are coded in GAMS using CPLEX 12.6.1 and implemented in the DTU HPC Cluster. The DTU HPC Cluster is a composite of a variety of hardware components of different technical characteristics. Therefore, we refer the reader to [30] for further and detailed information on the cluster and its components. We solve the raw HUC (without decomposition) and the SPDA2 using one node in a multi-threading configuration that counts on 10 Intel cores, while SPDA1 is implemented in a multi-node and multi-threaded environment. In particular, partitions are solved in parallel, each in a different node of the cluster with up to 10 Intel cores per node. Lastly, the reduced HUC problem is solved using again one node employing up to 10 Intel cores. In both SPDA1 and SPDA2 the subproblems are solved using the Gather-Update-Solve-Scatter Facility in GAMS.

Table II provides solutions times, achieved MIP gap (in percentage), number of scenarios retained in the final master problem, and expected and worst-case costs (in US dollars) for the three different solution strategies and for a different number of partitions. It is apparent that both SPDA1 and SPDA2 achieves remarkable time reductions without jeopardizing the solution quality at all. As expected, SPDA2 solves a last master problem with a few number of scenarios. However, in the intermediate steps of the algorithm, SPDA1 works with smaller master problems which may make it more computationally efficient for a low number of partitions. Note also that the time savings attained by the parallelization and decomposition algorithms gradually decrease with the number of partitions.

All the results obtained so far suggest that, in a practical setup, we could tackle the stochastic unit commitment problem as follows: We solve the proposed HUC formulation using the SPDA1 algorithm, starting with one single partition (corresponding to the most risk-averse solution) and gradually increasing the number of partitions until either the decision-maker is satisfied with the performance of the last solution obtained (recall that the level of conservatism of the solution decreases with the number of partitions) or the amount of available time has been reached. Bear in mind that the k-means and the hierarchical clustering methods with few partitions prompt solutions from the HUC formulation that are close to the stochastic one and that the SPDA1 is precisely most efficient when the number of partitions is small. Note that such a procedure is not that easy to implement with the method presented in [19], because the weighting factor α is a real number ranging from 0 to 1, both inclusive, and the relation between α and the efficient frontier can be non-linear and non-continuous. Therefore, it would be very difficult to define a step size for gradually increasing α as we can naturally do through the number of partitions.

TABLE II: Comparison of solving the HUC problem with and without decomposition.

		1 P (Robust)	3 P	5 P	8 P	10 P	50 P (Stochastic)
Time [min]	raw HUC	1322.7	687.6	767.4	721.6	788.1	
	SPDA1	28.6	32.7	64.8	102.8	107.7	140.2
	SPDA2		51.1	33.6	81.5	80.1	
GAP [%]	raw HUC		0.0299	0.0287	0.0297	0.0299	
	SPDA1	0.03	0.029	0.0297	0.0284	0.0206	0.0292
	SPDA2		0.03	0.0256	0.021	0.0284	
# Scen.	raw HUC	50	50	50	50	50	
	SPDA1	5	10	11	17	18	50
	SPDA2		8	9	12	14	
WC [10^4]	raw HUC		190.6821	190.6360	190.8254	190.7072	
	SPDA1	190.6035	190.6345	190.6426	190.6831	190.7932	191.8024
	SPDA2		190.7367	190.6573	190.6494	190.6978	
EC [10^4]	raw HUC		180.4438	180.4131	180.3176	180.3397	
	SPDA1	180.4169	180.4005	180.3980	180.3250	180.3448	180.2738
	SPDA2		180.4123	180.3832	180.3297	180.3562	

D. Comparison with a Scenario Reduction Technique.

We conclude our numerical study by comparing the unit-commitment solution given by SPDA1 or SPDA2 with that yielded by the widely-used fast forward scenario reduction technique [35]. The results of this comparison are summarized in Table III in terms of solution time (in seconds), MIP gap finally achieved (in percentage), and expected and worst-case costs (in US dollars). We use the fast forward selection algorithm to produce reduced scenario sets of varying cardinality. For the sake of a fair comparison, we make the cardinality of the reduced scenario set coincide with the number of scenarios retained (in the form of primal cuts) in the last master problem solved by SPDA1 or SPDA2 when one, two, five, eight and ten partitions are considered.

It is evident that, even though the use of a scenario reduction technique notably reduces the computational burden of the stochastic unit commitment problem, the unit-commitment solutions provided by SPDA1 and SPDA2 are significantly better in terms of the worst-case cost, while achieving expected costs close to that of the pure stochastic solution. Indeed, our solution approach yields unit-commitment plans that result in a worst-case cost that is, at least, 1.6% lower than the best worst-case cost provided by the scenario reduction technique (with 18 scenarios). Likewise, our solution approach provides unit-commitment plans that result in an expected cost that is, at most, 0.07% worse than that corresponding to the pure stochastic solution. This supports the conclusion that solving the proposed HUC formulation using either SPDA1 or SPDA2 constitutes a computationally efficient strategy to determine well-performing unit-commitment solutions in terms of the expected and worst-case costs.

TABLE III: Results relative to the applicaction of the fast forward scenario reduction technique [35].

# Scenarios	Time [min]	GAP [%]	EC [10^4]	WC [10^4]
5	2.8	0.0284	182.3564	212.0334
8	5.3	0.0293	181.0866	201.2821
9	7.4	0.0263	180.9487	200.6126
10	9.7	0.0294	180.9537	200.6188
11	11.1	0.0300	180.7298	197.9442
12	11.3	0.0291	180.6371	197.0820
14	11.7	0.0269	180.3185	194.0619
17	19.4	0.0287	180.2918	193.7436
18	25.6	0.0297	180.2912	193.7437

V. Conclusion and Future Research

In this paper we propose a new formulation of the unit commitment problem under uncertainty that allows us to find unit-commitment plans that perform relatively well in terms of both the expected and the worst-case system operating cost. The new formulation relies on clustering the scenario data set into a number of partitions. The expectation of the system operating cost is then taken over those scenarios that result in the worst-case dispatch cost within each partition. The conservatism of the so-obtained unit-commitment solution (that is, how close it is to the pure scenario-based stochastic or robust unit-commitment plan) is controlled via the user-specified number of partitions. We also develop parallelization-and-decomposition schemes to efficiently solve the proposed unit-commitment formulation. Our numerical results show that our scheme is able to dramatically reduce the required running time while improving the optimality of the solution found.

We envision three avenues of future research at least. First, we would like to investigate how our methodology adapts to the case where the partitions are continuous uncertainty sets. Second, it would also be interesting to explore different possibilities to speed up the SPDA algorithms, for example, by warm-starting the solution to the proposed HUC formulation for k partitions with the solution obtained for $k-1$ partitions. Finally, we would like to extend our formulation and the associated solution approach to a multi-stage setup.

References

[1] J. R. Birge and F. Louveaux, *Introduction to stochastic programming.* Springer Science & Business Media, 2011.

[2] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization.* Princeton University Press, 2009.

[3] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.

[4] F. Bouffard, F. D. Galiana, and A. J. Conejo, "Market-clearing with stochastic security-part i: formulation," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1818–1826, 2005.

[5] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1319–1327, 2008.

[6] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 900–910, 2009.

[7] A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit commitment for systems with significant wind penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 592–601, 2009.

[8] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 206–215, 2012.

[9] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

[10] L. Zhao and B. Zeng, "Robust unit commitment problem with demand response and wind energy," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, IEEE, 2012.

[11] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.

[12] H. Heitsch and W. Römisch, "Scenario reduction algorithms in stochastic programming," *Computational Optimization and Applications*, vol. 24, pp. 187-206, 2003.

[13] J. M. Morales, S. Pineda, A. J. Conejo, and M. Carrión, "Scenario Reduction for Futures Market Trading in Electricity Markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878–888, May 2009.

[14] S. Pineda and A. J. Conejo, "Scenario reduction for risk-averse electricity trading," in *IET Generation, Transmission & Distribution*, vol. 4, no. 6, pp. 694–705, 2010.

[15] L. Liu, G. Huang, Y. Liu, G. Fuller, and G. Zeng, "A fuzzy-stochastic robust programming model for regional air quality management under uncertainty," *Engineering Optimization*, vol. 35, no. 2, pp. 177–199, 2003.

[16] Y. Xu, G. Huang, X. Qin, and M. Cao, "SRCCP: a stochastic robust chance-constrained programming model for municipal solid waste management under uncertainty," *Resources, Conservation and Recycling*, vol. 53, no. 6, pp. 352–363, 2009.

[17] G. Liu, Y. Xu, and K. Tomsovic, "Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 227–237, 2016.

[18] B. Fanzeres, A. Street, and L. A. Barroso, "Contracting strategies for renewable generators: a hybrid stochastic and robust optimization approach," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1825–1837, 2015.

[19] C. Zhao and Y. Guan, "Unified stochastic and robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3353–3361, 2013.

[20] C. Zhao and Y. Guan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2587–2596, 2016.

[21] V. Gabrel and C. Murat and A. Thiele, "Recent advances in robust optimization: An overview," *European Journal of Operational Research*, vol. 235, no. 3, 2014.

[22] T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," in *The elements of statistical learning*, pp. 485–585, Springer, 2009.

[23] J. Paparrizos and L. Gravano, "k-shape: Efficient and accurate clustering of time series," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp. 1855–1870, ACM, 2015.

[24] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl, *et al.*, "Constrained k-means clustering with background knowledge," in *ICML*, vol. 1, pp. 577–584, 2001.

[25] B. F. Hobbs and Y. Ji, "Stochastic programming-based bounding of expected production costs for multiarea electric power system," *Operations Research*, vol. 47, no. 6, pp. 836–848, 1999.

[26] J.-P. Watson, F. D. Munoz, and B. Hobbs, "New bounding and decomposition approaches for MILP investment problems: Multi-area transmission and generation planning under policy constraints.," tech. rep., Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2014.

[27] A. Papavasiliou and S. S. Oren, "Multi-Area Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.

[28] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1371–1378, 2006.

[29] A. Papavasiliou, S. S. Oren, B. Rountree, "Applying High Performance Computing to Transmission-Constrained Stochastic Unit Commitment for Renewable Penetration," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1690–1701, 2015.

[30] "Central DTU HPC cluster —." http://www.cc.dtu.dk/?page_id=342. (Accessed on 03/21/2016).

[31] "Gams documentation center." http://www.gams.com/help/index. jsp?topic=\%2Fgams.doc\%2Fsolvers\%2Fguss\%2Findex.html& anchor=GUSS_APP_DEA_MODELING. (Accessed on 06/01/2016).

[32] "Power Systems Test Case Archive - UWEE." https://www.ee. washington.edu/research/pstca/. (Accessed on 04/19/2016).

[33] "The Real Lab - Renewable Energy Analysis Laboratory." http://www.ee.washington.edu/research/real/library.html. (Accessed on 03/21/2016).

[34] P. Pinson *et al.*, "Wind energy: Forecasting challenges for its operational management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.

[35] H. Heitsch and W. Römisch, "Scenario reduction algorithms in stochastic programming," *Computational optimization and applications*, vol. 24, no. 2-3, pp. 187–206, 2003.

APPENDIX

This Appendix provides the wind power production scenarios used in the case study based on the IEEE 14-node system (Table IV) and the location of wind farms in the 3-Area RTS-96 system (Table V).

TABLE IV: Wind power production scenarios [MW] for the IEEE 14-node system.

	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10
t1	66.45	67.06	65.39	57.12	42.9	37.8	38.64	29.12	27.96	22.35
t2	107.4	102.34	86.71	85.2	70.29	65.79	59.2	51.52	40.26	35.6
t3	113.25	103.32	100.23	90.6	74.47	66.24	63.76	51.45	47.82	37.25
t4	127.5	102.2	113.1	100.56	90.2	78.57	61.04	58.03	46.5	40.85
t5	124.2	113.68	106.08	109.2	94.6	78.48	68.32	55.3	51.84	37.95
t6	128.25	113.4	110.63	88.92	93.72	78.57	68.8	55.86	50.22	41.75
t7	152.1	147.98	127.14	116.04	101.31	91.71	78.48	73.85	60.78	52.65
t8	158.55	139.16	117.65	120.36	109.89	80.82	73.68	69.09	63.48	53.7
t9	120.15	119.14	107.77	93.24	81.4	70.65	58.64	51.24	46.2	42.5
t10	85.8	86.94	81.9	81.6	72.93	60.03	47.44	43.05	37.98	32.55
t11	148.2	141.68	133.51	128.88	105.05	89.37	82.16	67.41	52.74	52.7
t12	131.25	129.08	113.75	104.04	98.01	86.94	69.6	68.46	61.38	45.6
t13	135.45	120.12	117.65	102.36	93.61	71.28	62.4	59.5	46.98	39.55
t14	118.5	109.9	112.45	96	81.62	76.14	59.36	52.36	51.48	41.2
t15	110.25	114.52	104.39	90.84	92.51	72.18	62.4	50.96	47.46	39.9
t16	44.4	43.4	41.21	35.04	35.09	29.79	22.08	21.84	22.26	14.85
t17	5.55	6.02	5.85	4.44	4.51	3.87	3.44	2.8	2.22	2.05
t18	14.25	11.2	10.79	8.76	8.25	6.12	6.96	5.39	4.74	3.8
t19	17.1	13.72	14.69	12	10.45	8.73	7.04	8.26	6.06	4.65
t20	7.95	6.44	6.89	6.72	6.16	3.87	3.6	3.92	3.36	2.5
t21	9.6	7.84	7.41	6.84	7.7	5.67	5.44	3.92	3.78	3.05
t22	87.75	70.98	67.6	62.52	64.46	51.84	45.76	38.29	34.74	27.85
t23	119.4	104.58	114.14	98.16	88.33	70.47	66.72	61.88	49.98	45.15
t24	82.65	69.58	62.53	63.48	57.09	44.73	43.6	39.69	33.6	26.65

TABLE V: Location of wind farms in the IEEE 3-Area RTS-96 system.

Unit	Node	Unit	Node	Unit	Node	Unit	Node	Unit	Node
f_1	103	f_4	121	f_7	216	f_{10}	303	f_{13}	316
f_2	105	f_5	203	f_8	221	f_{11}	305	f_{14}	321
f_3	116	f_6	205	f_9	223	f_{12}	307	f_{15}	323

Ignacio Blanco received the B.S. degree in Energy Engineering in 2013 from the Rey Juan Carlos University, Spain. In 2015 he obtained the M.Sc. degree in Energy Management from the Technical University of Madrid. He is currently completing his Ph.D. at the Technical University of Denmark.

His research interests are large-scale optimization, data analysis, energy planning and decision-making under uncertainty for integrated energy systems.

Juan M. Morales (S'07-M'11-SM'16) received the Ingeniero Industrial degree from the Universidad de Málaga, Málaga, Spain, in 2006, and a Ph.D. degree in Electrical Engineering from the Universidad de Castilla-La Mancha, Ciudad Real, Spain, in 2010. He is currently an associate professor in the Department of Applied Mathematics at the University of Málaga in Spain.

His research interests are in the fields of power systems economics, operations and planning; energy analytics and smart grids; decision-making under under uncertainty, and electricity markets.