

Temporal Decomposition-Based Stochastic Economic Dispatch for Smart Grid Energy Management

Farnaz Safdarian¹, *Student Member, IEEE*, and Amin Kargarian¹, *Member, IEEE*

Abstract—This paper presents a temporal decomposition strategy to decompose security-constrained economic dispatch (SCED) over the scheduling horizon with the goal of reducing its computational burden and enhancing its scalability. A set of subproblems, each with respect to demand response, normal constraints, and $N - 1$ contingency corrective actions at a subhorizon, is formulated. The proposed decomposition deals with computational complexities originated from intertemporal interdependencies of system equipment, i.e., generators' ramp constraints and state of charge of storage devices. The concept of overlapping intervals is introduced to make SCED subproblems solvable in parallel. Intertemporal connectivity related to energy storage is also modeled in the context of temporal decomposition. Besides, reserve up and down requirements are formulated as data-driven nonparametric chance constraints to account for wind generation uncertainties. The concept of ϕ -divergence is used to convert nonparametric chance constraints to more conservative parametric constraints. A reduced risk level is calculated with respect to wind generation prediction errors to ensure the satisfaction of system constraints with a confidence level after the true realization of uncertainty. Auxiliary problem principle is applied to coordinate SCED subproblems in parallel. Numerical results on three test systems show the effectiveness of the proposed algorithm.

Index Terms—Temporal decomposition, distributed optimization, economic dispatch, demand response, energy storage, chance constraint.

NOMENCLATURE

Indices, Sets, and Parameters

c	Index for contingencies.
i, j	Index for buses.
ij	Index for lines.
k	Index for iterations.
m	Index for subproblems and subhorizons.
s	Index for storage devices.
t	Index for time intervals.

Manuscript received September 19, 2019; revised January 23, 2020 and April 7, 2020; accepted May 6, 2020. Date of publication May 12, 2020; date of current version August 21, 2020. This work was supported in part by the National Science Foundation under Grant ECCS-1944752, and in part by Louisiana Board of Regents under Grant LEQSF (2016-19)-RD-A-10. Paper no. TSG-01390-2019. (*Corresponding author: Amin Kargarian.*)

The authors are with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803 USA (e-mail: fsafda1@lsu.edu; kargarian@lsu.edu).

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

Digital Object Identifier 10.1109/TSG.2020.2993781

u	Index for units.
w	Index for wind farms.
a, b, o	Cost coefficients for generating units.
b_{sh}	load shedding cost.
$PD_{i,t}$	Power demand at bus i at time t .
PD_t	Total power demand at time t .
SP_m	Subproblem m corresponding to subhorizon m .
SP_m^+	Subproblem m with dummy time intervals.
$ICC_{m,m+1}$	Intertemporal consistency constraints between two consecutive subproblems m and $m + 1$.
UR_u	Ramping up limit of unit u .
DR_u	Ramping down limit of unit u .
T	Overall time horizon.
Ξ_{m+1}^{*k-1}	Values of shared variables determined by subproblems SP_{m+1}^+ at iteration $k - 1$.
λ	Vector of Lagrange multipliers.
β, γ	Tuning parameters.
ρ	Penalty factor.
$\underline{z}, \bar{z}$	Minimum and maximum of a variable z .
μ	Efficiency of storage.

Variables

$E_{s,t}$	State of charge of storage s at time t .
$I_{ch,s}, I_{dc,s}$	Binary variables for modeling storage charging and discharging modes.
$p_{u,t}$	Generation of unit u at time t .
$p_{ch,s,t}$	Charging power of storage s at time t .
$p_{dc,s,t}$	Discharging power of storage s at time t .
$p_{ij,t}$	Flow in line ij at time t .
$p_{sh,i,t}$	Load shedding at bus i at time t .
$\delta_{i,t}$	Voltage angle at bus i at time t .
$r_{u,t}^d$	Reserve down provided by unit u at time t .
$r_{u,t}^u$	Reserve up provided by unit u at time t .
$\Xi_{m-1,m}$	Set of shared variables pertaining to overlapping time intervals between subproblems $m - 1$ and m .
$\Xi_m^k, \Xi_m^{k'}$	Shared variables of subproblem m with its previous and next neighbors at iteration k .
z_t^c	Values of a variable z at time t after contingency c .

Uncertainty Related Parameters and Functions

$f_{w,t}(\cdot)$	True wind power PDF at interval t .
$\hat{f}_{w,t}(\cdot)$	Estimated wind power PDF at interval t .

$E(\tilde{P}_{w,t})$	Expected value of wind power generation at time t .
$\tilde{P}_{w,t}$	Wind power generation at time t .
ξ	A random variable.
α	Risk level of chance constraints.
$\mathbb{P}$	Probability measure.
$\hat{\mathbb{P}}$	Estimated distribution.
$\widehat{CDF}^{-1}$	Estimated quantile function (i.e., inverse CDF).

I. INTRODUCTION

INTEGRATING renewable energy sources, storage devices, and demand response into power systems have brought new challenges to system operation and planning problems, such as security-constrained economic dispatch (SCED) [1]. Two main challenges include effective uncertainty modeling and solving large optimization problems. The size of SCED grows by increasing the number of uncertainty sources, the size of the system, and the number of considered contingencies. Another important factor that increases the size and computational burden of SCED drastically is the number of time intervals of the considered scheduling horizon. Operators should solve SCED with different time horizons, e.g., one day or one week, for various power system analysis purposes. Intertemporal constraints, e.g., generators ramp limits and state of charge of storage devices, increase the complexity and solution time of SCED, particularly if the number of time intervals increases.

Decomposition approaches are presented in the literature to address this challenge and solve SCED in a reasonable time span. The majority of existing papers deal with decomposing SCED over geographical areas [2]–[5]. In geographical decomposition, a power system is decomposed into several smaller zones and a local subproblem is formulated for each zone. References [2], [6], [7] review decomposition and distributed optimization algorithms, such as alternating direction method of multipliers [8], analytical target cascading [9], optimality condition decomposition [10], and auxiliary problem principle (APP) [11], that can be used to solve SCED and optimal power flow. In [12], a distributed bisection algorithm is proposed for economic dispatch to minimize the aggregated cost of a network. In [13], a decentralized and self-organizing solution framework for economic dispatch is proposed. In [14], the distributed economic dispatch and demand response initiatives for grid-connected microgrids with high-penetration of wind power are studied. Distributed transmission+distribution SCED using multi-parametric quadratic programming is presented in [3]. Reference [15] introduces a consensus-based control scheme to solve economic dispatch distributedly. Decentralized approaches for solving economic dispatch in smart grids are presented in [16]–[20].

However, the computational complexity originating from intertemporal connectivity between constraints is not considered in geographical decomposition strategies. If the number of time intervals grows, the computational burden of SCED increases. In addition, the solution time increases drastically with considering $N - 1$ security criteria, which contribute to increasing the number of intertemporal constraints. To solve multi-interval problems, such as look-ahead SCED [21]–[23],

centralized optimization methods may face computational challenges as the size or the number of considered intervals grows. Geographical decompositions would reduce the solution time by breaking the problem, e.g., look-ahead SCED, over geographical areas, but they do not deal with intertemporal constraints.

On the other hand, while some distributed approaches consider uncertainties [24], many others ignore uncertainties as they complicate distributed optimization. Chance-constrained programming is an efficient method for modeling uncertainties in short-term operation problems, such as optimal power flow, SCED, and unit commitment. Chance constraints do not hinder the complexity of distributed optimization significantly. The majority of existing papers, particularly for distributed SCED, use parametric chance-constrained methods [25], [26] where a known probability distribution function (PDF) is considered for uncertain parameters. However, a random parameter may not belong to any class of known distribution function. Parameters and shape of PDFs might change depending on the weather condition and geographical location.

In this paper, we aim to decrease the solution time of the SCED problem taking into account wind generation uncertainties, controllable loads, energy storage constraints, and $N - 1$ contingency corrective actions. A temporal decomposition strategy is proposed to decompose the problem over the time horizon and create a set of subhorizons, each representing a segment of the considered scheduling horizon. The concept of overlapping time intervals is introduced to model intertemporal connectivity corresponding to generators' ramp limitations and energy storage constraints for transition between two consecutive subhorizons and in order to make SCED subproblems solvable in parallel. The proposed decomposition strategy relieves complexities originated from intertemporal constraints and can be considered as the complement of geographical decomposition. That is, temporal decomposition can be combined with geographical decomposition to further reduce the SCED solution time. Wind generation uncertainties are considered and data-driven nonparametric chance constraints are presented to model reserve up and down requirements. The concept of ϕ -divergence is used to calculate a reduced risk level with respect to wind forecast errors. Nonparametric chance constraints are reformulated with their more conservative, equivalent parametric constraints to ensure the satisfaction of system constraints after the true realization of uncertainty. APP, along with a suitable initialization technique, is applied to coordinate SCED subproblems in parallel. To evaluate the effectiveness of the distributed approach, we solve a week-ahead SCED for a 6-bus system, the IEEE 24-bus system, and a 472-bus system. Promising results are obtained as compared to those of the conventional centralized algorithm.

The remainder of this paper is organized as follows. Section II presents temporal decomposition and data-driven chance constraints. In Section III, modeling interdependencies between consecutive subhorizons is explained. A coordination strategy is discussed in Section IV. Simulation results are illustrated in Section V, and concluding remarks are provided in Section VI.

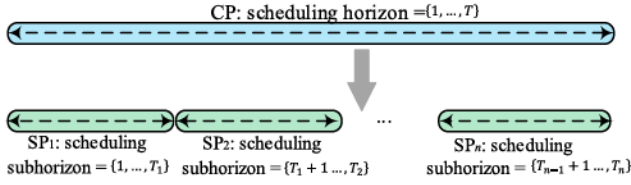


Fig. 1. Time decomposition concept.

II. TEMPORAL DECOMPOSITION OF SCED

Consider a centralized SCED problem (CP) with an overall time horizon of T . As shown in Fig. 1, we propose to decompose CP over the time horizon to create n SCED subproblems (SPs), each including a subhorizon of T so that:

$$\{1, \dots, T_1\} \cup \dots \cup \{T_{n-1} + 1, \dots, T_n\} = \{1, \dots, T\} \quad (1)$$

Each SP contains fewer variables and constraints, especially intertemporal constraints, than the original CP. Hence, subproblems are computationally less expensive than CP.

A. Deterministic Constraints of Subproblem m

Consider subhorizon m with the scheduling interval of $\{T_{m-1} + 1, \dots, T_m\} \subseteq \{1, \dots, T\}$. A single-stage nonparametric chance-constrained-based SCED is formulated for SP_m whose objective function is to minimize operational costs over subhorizon m . The first three terms of (2) are generators cost function, the fourth term is the reserve cost, and the fifth term is the load shedding cost [1].

$$\begin{aligned} \min \quad & \sum_t \sum_u a_u p_{u,t}^2 + b_u p_{u,t} + o_{u,t} + b_u r_{u,t}^{up} \\ & + \sum_i b_{sh,i,t} p_{sh,i,t} \end{aligned} \quad (2)$$

The continuous decision variables are $\{E_{s,t}, p_{u,t}, p_{ch,s,t}, p_{dc,s,t}, p_{ij,t}, p_{sh,i,t}, \delta_{i,t}, r_{u,t}^d, r_{u,t}^{up}\}$ and the binary decision variables are $\{I_{ch,s}, I_{dc,s}\}$ under normal conditions and each contingency c . Constraints of SP_m include system and component restrictions under normal condition and $N - 1$ security criteria $\forall t \in \{T_{m-1} + 1, \dots, T_m\}$. Constraints under normal condition are power balance equalities (3)-(4) and transmission line limitation (6), load shedding limitation (7), upper and lower bounds of generating units (8)-(9), units ramping restriction (10)-(11), units spinning reserve up and down constraints (12)-(15), storage state of change at time t (16), storage charge and discharge power restrictions (17)-(18), constraints to avoid simultaneous storage charge and discharge (19), and upper and lower bounds of storage energy level (20) [1], [27], [28]. To ensure system security after the occurrence of a contingency, corrective actions should be scheduled by adjusting control variables (e.g., power generated by units and storage power charge/discharge) within a reasonable range. A set of new variables are defined for each contingency c . System and equipment constraints corresponding to each contingency c , known as $N - 1$ security criteria [1], [27], [28], include power balance constraints (21)-(22), line flow limits (24), load shedding constraints (25)-(26), thermal units upper and lower bounds considering reserve values (27)-(32),

thermal units adjustment capabilities after the occurrence of a contingency as compared to their generation under normal condition (33), and storage limitations after the occurrence of a contingency (34)-(38). Expression (34) shows the amount of energy that a storage unit must provide to help the system after the occurrence of a contingency [28]. After an outage, fast-response storage units inject or extract power instantly to bring line flows back down within their short-term emergency rating. The power injections or extractions from the storage units remain constant for a period of τ_1 (e.g., 5 minutes) until generators start ramping. During the ramping period τ_2 (e.g., 10 minutes), storage units reduce their injections or extractions until they reach zero, while generators ramp their output up or down. Flows in overloaded lines decrease until they reach their long-term emergency rating [28].

$$\begin{aligned} \sum_u p_{u,t} + \sum_s (p_{dc,s,t} - p_{ch,s,t}) + \sum_w E(\tilde{P}_{w,t}) \\ = \sum_i PD_{i,t} - p_{sh,i,t} \quad \forall t \end{aligned} \quad (3)$$

$$p_{i,t} - PD_{i,t} = \sum_j \frac{\delta_{i,t} - \delta_{j,t}}{X_{ij}} \quad \forall i, \quad \forall t \quad (4)$$

$$\delta_{ref,t} = 0 \quad \forall t \quad (5)$$

$$\underline{P}_{ij} \leq p_{ij,t} = \frac{\delta_{i,t} - \delta_{j,t}}{X_{ij}} \leq \bar{P}_{ij} \quad \forall ij, \quad \forall t \quad (6)$$

$$\underline{P}_{sh,i} \leq p_{sh,i,t} \leq \bar{P}_{sh,i} \quad \forall i, \quad \forall t \quad (7)$$

$$\underline{P}_{u,t} \leq p_{u,t} + r_{u,t}^{up} \leq \bar{P}_{u,t} \quad \forall u, \quad \forall t \quad (8)$$

$$\underline{P}_{u,t} \leq p_{u,t} - r_{u,t}^d \leq \bar{P}_{u,t} \quad \forall u, \quad \forall t \quad (9)$$

$$(p_{u,t} + r_{u,t}^{up}) - (p_{u,t-1} - r_{u,t}^d) \leq UR_u \quad \forall u, \quad \forall t \quad (10)$$

$$(p_{u,t-1} + r_{u,t-1}^{up}) - (p_{u,t} - r_{u,t}^d) \leq DR_u \quad \forall u, \quad \forall t \quad (11)$$

$$0 \leq r_{u,t}^{up} \leq UR_{u,10min} \quad \forall u, \quad \forall t \quad (12)$$

$$0 \leq r_{u,t}^{up} \leq \bar{P}_{u,t} - p_{u,t} \quad \forall u, \quad \forall t \quad (13)$$

$$0 \leq r_{u,t}^d \leq DR_{u,10min} \quad \forall u, \quad \forall t \quad (14)$$

$$0 \leq r_{u,t}^d \leq p_{u,t} - \underline{P}_{u,t} \quad \forall u, \quad \forall t \quad (15)$$

$$E_{s,t} = E_{s,t-1} + \left(\mu p_{ch,s,t} - \frac{p_{dc,s,t}}{\mu} \right) \Delta t \quad \forall s, \quad \forall t \quad (16)$$

$$I_{ch,s,t} \cdot \underline{P}_{ch,s} \leq p_{ch,s,t} \leq I_{ch,s,t} \cdot \bar{P}_{ch,s} \quad \forall s, \quad \forall t \quad (17)$$

$$I_{dc,s,t} \cdot \underline{P}_{dc,s} \leq p_{dc,s,t} \leq I_{dc,s,t} \cdot \bar{P}_{dc,s} \quad \forall s, \quad \forall t \quad (18)$$

$$I_{dc,s,t} + I_{ch,s,t} \leq 1 \quad \forall s, \quad \forall t \quad (19)$$

$$\underline{E}_s \leq E_{s,t} \leq \bar{E}_s \quad \forall s, \quad \forall t \quad (20)$$

$$\begin{aligned} \sum_u p_{u,t}^c + \sum_s (p_{dc,s,t}^c - p_{ch,s,t}^c) + \sum_w E(\tilde{P}_{w,t}^c) \\ = \sum_i D_{i,t} - p_{sh,i,t}^c \quad \forall t, \quad \forall c \end{aligned} \quad (21)$$

$$p_{i,t}^c - D_{i,t} = \sum_j \frac{\delta_{i,t}^c - \delta_{j,t}^c}{X_{ij}} \quad \forall i, \quad \forall t, \quad \forall c \quad (22)$$

$$\delta_{ref,t}^c = 0 \quad \forall t, \quad \forall c \quad (23)$$

$$\underline{P}_{ij}^c \leq p_{ij,t}^c = \frac{\delta_{i,t}^c - \delta_{j,t}^c}{X_{ij}} \leq \bar{P}_{ij}^c \quad \forall ij, \quad \forall c \quad (24)$$

$$\underline{P}_{sh,i}^c \leq p_{sh,i,t}^c \leq \bar{P}_{sh,i}^c \quad \forall i, \quad \forall t, \quad \forall c \quad (25)$$

$$|p_{sh,i,t} - p_{sh,i,t}^c| \leq \Delta \quad \forall u, \quad \forall t, \quad \forall c \quad (26)$$

$$\underline{p}_{u,t} \leq p_{u,t}^c + r_{u,t}^{up,c} \leq \bar{p}_{u,t} \quad \forall u, \quad \forall t, \quad \forall c \quad (27)$$

$$\underline{p}_{u,t} \leq p_{u,t}^c - r_{u,t}^{d,c} \leq \bar{p}_{u,t} \quad \forall u, \quad \forall t, \quad \forall c \quad (28)$$

$$0 \leq r_{u,t}^{up,c} \leq UR_{u,10min} \quad \forall u, \quad \forall t, \quad \forall c \quad (29)$$

$$0 \leq r_{u,t}^{up,c} \leq \bar{p}_{u,t} - p_{u,t} \quad \forall u, \quad \forall t, \quad \forall c \quad (30)$$

$$0 \leq r_{u,t}^{d,c} \leq DR_{u,10min} \quad \forall u, \quad \forall t, \quad \forall c \quad (31)$$

$$0 \leq r_{u,t}^{d,c} \leq p_{u,t} - \underline{p}_{u,t} \quad \forall u, \quad \forall t, \quad \forall c \quad (32)$$

$$|p_{u,t} - p_{u,t}^c| \leq \Delta \quad \forall u, \quad \forall t, \quad \forall c \quad (33)$$

$$E_{s,t}^c = E_{s,t} + (\tau_1 + 0.5\tau_2) \left(\mu p_{ch,s,t}^c - \frac{p_{dc,s,t}^c}{\mu} \right) \Delta t \quad \forall s, \quad \forall t, \quad \forall c \quad (34)$$

$$I_{ch,s,t}^c \cdot \underline{p}_{ch,s} \leq p_{ch,s,t}^c \leq I_{ch,s,t}^c \cdot \bar{p}_{ch,s} \quad \forall s, \quad \forall t, \quad \forall c \quad (35)$$

$$I_{dc,s,t}^c \cdot \underline{p}_{dc,s} \leq p_{dc,s,t}^c \leq I_{dc,s,t}^c \cdot \bar{p}_{dc,s} \quad \forall s, \quad \forall t, \quad \forall c \quad (36)$$

$$I_{dc,s,t}^c + I_{ch,s,t}^c \leq 1 \quad \forall s, \quad \forall t, \quad \forall c \quad (37)$$

$$\underline{E}_s \leq E_{s,t}^c \leq \bar{E}_s \quad \forall s, \quad \forall t, \quad \forall c. \quad (38)$$

B. Probabilistic Constraints of Subproblem m

In addition to (3)-(38), adequate reserve up and down should be provided to compensate wind generation forecast errors. Reserve requirements are formulated as probabilistic constraints. Chance-constrained programming is a suitable approach for modeling reserve requirements [29]. Wind power depends on different factors and may not follow any known class of density functions. Chance constraints are sensitive to PDFs. If the true realizations of wind generation do not match with the assumed PDF, the probability of satisfaction of parametric chance constraints may violate the predetermined confidence level that results in system security degradation.

We formulate two data-driven nonparametric chance constraints (41) and (42) to ensure that generation and reserve down/up provided by thermal units and storage satisfy load with a confidence level of $1 - \alpha$ if wind power goes above or falls below its forecast values. A confidence set for each interval t is specified by the ϕ -divergence function and tolerance $d_{w,t}$ representing the size of the confidence set [30]. We can use the worst distribution in the confidence set (i.e., the PDF with the largest distance from the true PDF within a tolerance d) to formulate nonparametric chance constraints (41) and (42). Using historical data, we define the confidence set for wind generation at time t as [30]:

$$\mathcal{D}_{w,t} = \left\{ \mathbb{P}_{w,t} \in \mathcal{M}_+ : D_{w,t}(\mathbb{P}_{w,t} \| \hat{f}_{w,t}) \leq d_{w,t}, f_{w,t} = \frac{d\mathbb{P}_{w,t}}{d\tilde{\xi}_t} \right\} \quad (39)$$

where $\mathbb{P}_{w,t}$ represents the ambiguous true distribution function for wind power at interval t , and $\mathcal{M}_+$ is the set of all distribution functions. The distance between the estimated and true wind power PDFs, or ϕ -divergence $D_{s,t}(\cdot \| \cdot)$, is defined as [30]–[32]:

$$D_{w,t}(\mathbb{P}_{w,t} \| \hat{f}_{w,t}) = \int_{\mathbb{R}^m} \phi \left(\frac{f_{w,t}(\tilde{\xi}_t)}{\hat{f}_{w,t}(\tilde{\xi}_t)} \right) \hat{f}_{w,t}(\tilde{\xi}_t) d\tilde{\xi}_t \quad (40)$$

where $\phi(\cdot)$ is a convex function on $\mathbb{R}^+$. The summation of generation minus reserve down should be less than or equal to demand. The reserve down is required for situations in which the true realization of wind generation is larger than its predicted value.

$$\inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P} \left\{ \sum_u p_{u,t} - r_{u,t}^d + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} + \bar{p}_{w,t} \leq PD_t \right\} \geq 1 - \alpha \quad \forall t \quad (41)$$

where the $\inf$ operator represents PDF with the largest distance from the true PDF within a tolerance d . The nonparametric reserve up chance constraints is required when the true realization of wind generation is less than its expected value.

$$\inf_{\mathbb{P} \in \mathcal{D}_{w,t}} \mathbb{P} \left\{ \sum_u p_{u,t} + r_{u,t}^{up} + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} + \bar{p}_{w,t} \geq PD_t \right\} \geq 1 - \alpha \quad \forall t \quad (42)$$

We replace the risk level α with a reduced nonnegative risk level α'_+ with respect to the divergence function $\phi(\cdot)$ and the divergence tolerance d . Then, as proven in [31], the predicted PDF for wind generation, $\hat{\mathbb{P}}$, can be used to reformulate (41) and (42) as parametric chance constraint using $\hat{\mathbb{P}}$ instead of $\inf_{\mathbb{P} \in \mathcal{D}}$. This procedure is described below.

- Impose no assumption on the PDF of wind generation at each time interval, and estimate the unknown PDF ($\hat{\mathbb{P}}$) of random parameters from historical data using adaptive kernel density estimator (AKDE) [33].
- Form a histogram set of real data, from historical wind generation at time t . Determine pointwise errors between the histogram and $\hat{\mathbb{P}}$, and calculate the distance d from the square of errors as [30]:

$$d = SE_{1-\alpha}^2. \quad (43)$$

- Choose an appropriate divergence function $\phi(\cdot)$, such as the χ divergence of order two that is suitable for small risk levels, and solve a univariate optimization problem to find α' for each time interval t as [31]:

$$\alpha' = \alpha - \frac{\sqrt{d^2 - 4d(\alpha - \alpha^2)} - (1 - 2\alpha)d}{2d + 2}. \quad (44)$$

- Avoid negative risk levels as:

$$\alpha'^+ = \max(0, \alpha'). \quad (45)$$

- In (41) and (42), replace $\inf_{\mathbb{P} \in \mathcal{D}}$ by $\hat{\mathbb{P}}$ and α by α'_+ . We now rewrite reserve down constraint (41) as follows:

$$\hat{\mathbb{P}} \left\{ \bar{p}_{w,t} \leq PD_t - \left(\sum_u p_{u,t} - r_{u,t}^d + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} \right) \right\} \geq 1 - \alpha'_+ \quad \forall t \quad (46)$$

Knowing that in the probability theory $\mathbb{P}\{Y \leq y\} = 1 - \mathbb{P}\{Y > y\}$, we reorganize the reserve up (45) as:

$$\hat{\mathbb{P}} \left\{ \tilde{P}_{w,t} \leq PD_t - \left(\sum_u p_{u,t} + r_{u,t}^{up} + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} \right) \right\} \leq \alpha_t^{'+} \quad \forall t \quad (47)$$

The left sides of (46) and (47) are the estimated cumulative distribution function (CDF) of wind generation $\tilde{P}_{w,t}$ at time t . By taking the CDF inverse, (46) and (47) are expressed as:

$$PD_t - \left(\sum_u p_{u,t} + r_{u,t}^{up} + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} \right) \leq \widehat{CDF}_{\tilde{P}_{w,t}}^{-1}(\alpha_t^{'+}) \quad \forall t \quad (48)$$

$$PD_t - \left(\sum_u p_{u,t} - r_{u,t}^d + \sum_s p_{dc,s,t} - p_{ch,s,t} + \sum_i p_{sh,i,t} \right) \geq \widehat{CDF}_{\tilde{P}_{w,t}}^{-1}(1 - \alpha_t^{'+}) \quad \forall t \quad (49)$$

The left-hand sides of (48) and (49) are variables and the right-hand sides are constant values. These linear constraints, which are linear equivalents of (41) and (42), ensure that the reserve up and down constraints are satisfied for all wind power distribution functions in the confidence set.

The reduced risk level makes chance constraints more conservative to account for wind power PDFs estimation errors. The level of conservativeness or the reduced risk level depends on errors of estimated PDFs. If more data or better estimation approaches are available to obtain more accurate wind power PDFs, the reduced risk level becomes closer to the user's predetermined risk level and the level of conservativeness of constraints goes down.

III. MODELING SUBHORIZONS TEMPORAL INTERDEPENDENCIES

A. Recovering Centralized SCED From Subproblems

Consider two consecutive subhorizons m and $m+1$ that cover intervals $\{T_{m-1} + 1, \dots, T_m\}$ and $\{T_m + 1, \dots, T_{m+1}\}$, respectively. Formulating SP_m and SP_{m+1} as presented in the previous section models intertemporal constraints inside each subhorizon, but it disregards constraints for transition between two consecutive subhorizons. Thus, the constraint sets of subproblems (denoted by CSP) are mutually exclusive and their intersection is empty.

$$CSP_m \cap CSP_{m+1} = \emptyset \quad (50)$$

This is not correct from the perspective of the centralized SCED. Although the overall objective function of the centralized SCED can be recovered by combining all subproblems' local objective functions, its constraint set cannot be reconstructed since in the centralized SCED we have:

$$CSP_{T_m} \cap CSP_{T_{m+1}} \neq \emptyset \quad (51)$$

That means collecting all elements (i.e., objective terms and constraints sets) of subproblems by taking their union does not

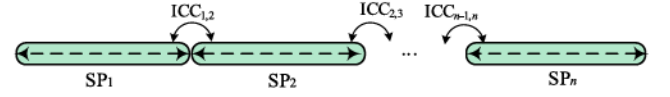


Fig. 2. ICCs for modeling transition between subproblems.

construct the centralized problem.

$$CP \neq SP_1 \cup SP_2 \dots \cup SP_m \cup SP_{m+1} \cup \dots \cup SP_n \quad (52)$$

Hence, solving these independent subproblems provide a solution that may be infeasible or suboptimal from the perspective of the whole scheduling horizon T . To solve this challenge, as shown in Fig. 2, a set of intertemporal consistency constraints (ICCs) must be added to the problem to model transition between subproblems. Now, the centralized SCED can be reconstructed by collecting all elements of SP_1 to SP_n and ICCs.

$$CP = SP_1 \cup ICC_{1,2} \cup SP_2 \cup ICC_{2,3} \cup \dots \cup ICC_{n-1,n} \cup SP_n \quad (53)$$

Although these ICCs are required for accurate SCED decomposition over the time horizon, the two following challenges must be addressed.

1. How to model ICCs knowing that there are several intertemporal constraints corresponding to system requirements and equipment models.
2. ICCs are obstacles for solving subproblems in a distributed manner independently as $CSP_{T_m} \cap CSP_{T_{m+1}} = ICC_{T_m, T_{m+1}}$.

B. Modeling Interdependencies With Overlapping Intervals

Intertemporal constraints of SCED include generators' ramp limits (10)-(11), and the state of charge of energy storage (16) and (34). Modeling these constraints as ICCs in the context of temporal decomposition becomes more challenging, taking into account possible corrective actions after the occurrence of a contingency, reserve up/down requirements, wind uncertainties, and demand response. A naïve approach is to start solving the first subproblem and fix the state of the last interval of a subproblem in the first interval of its next subproblem. SP_1 is solved first and the variable values at the last interval are sent to SP_2 . Ramping and storage constraints in SP_2 are formulated using these values as a fixed initial state. This procedure is carried out until SP_n is solved. This process is neither sequential nor iterative. The flow of information exchange is from one subproblem to its next subproblems. No feedback is sent from one subproblem to its previous one. Although this makes the solution feasible, the first few subproblems force their desired states to other subproblems as hard constraints that lead to a suboptimal solution for the overall SCED problem.

To ensure both feasibility and optimality of the obtained solution by the temporal decomposition strategy, we propose the concept of *overlapping time intervals* to facilitate modeling ICCs between subproblems. We add dummy intervals to the ending of all subproblems, except for the last SP_n . The dummy time interval are copies of the first intervals of all subproblems, except for SP_1 . Let us name subproblems with dummy

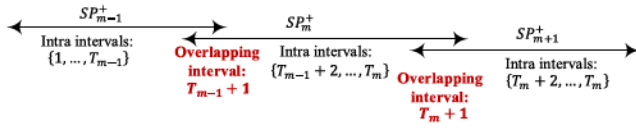


Fig. 3. Three consecutive subhorizons with overlapping time intervals.

time intervals as SP^+ . Consider Fig. 3 that shows a horizon that is decomposed into three consecutive subproblems SP_{m-1}^+ , SP_m^+ and SP_{m+1}^+ . The intra intervals of SP_{m-1}^+ , SP_m^+ and SP_{m+1}^+ are $\{1, \dots, T_{m-1}\}$, $\{T_{m-1} + 2, \dots, T_m\}$, $\{T_m + 2, \dots, T_{m+1}\}$, respectively. The overlapping time intervals, indicated by red in Fig. 3, are $T_{m-1} + 1$ between SP_{m-1}^+ , SP_m^+ , and $T_m + 1$ between SP_m^+ and SP_{m+1}^+ . Since each overlapping time interval is considered in two neighboring subhorizons, internal intertemporal constraints of SP_{m-1}^+ , SP_m^+ and SP_{m+1}^+ will be satisfied locally with respect to constraints of overlapping intervals. Now, subproblems are rewritten as follows:

$$\begin{aligned}
 &SP_{m-1}^+: \min \quad (2) \\
 &\quad s.t. \quad (3) - (38), (48) \& (49) \\
 &\quad \quad \forall t \in \{1, 2, \dots, T_{m-1}\} \cup \{T_{m-1} + 1\} \\
 &SP_m^+: \min \quad (2) \\
 &\quad s.t. \quad (3) - (38), (48) \& (49) \\
 &\quad \quad \forall t \in \{T_{m-1} + 1\} \cup \{T_{m-1} + 2, \dots, T_m\} \cup \{T_m + 1\} \\
 &SP_{m+1}^+: \min \quad (2) \\
 &\quad s.t. \quad (3) - (38), (48) \& (49) \\
 &\quad \quad \forall t \in \{T_m + 1\} \cup \{T_m + 2, \dots, T_{m+1}\}
 \end{aligned}$$

Variables and constraints of each overlapping time interval appear in two neighboring subproblems. By collecting all elements of subproblems, the intersection of consecutive subproblems, which includes variables, objective terms, and constraints at overlapping intervals, is counted twice. The centralized problem can be recovered by taking the union of SP_{m-1}^+ , SP_m^+ and SP_{m+1}^+ and subtracting the intersection of each two consecutive subproblems from it as follows:

$$\begin{aligned}
 CP &= (SP_{m-1}^+ \cup SP_m^+ \cup SP_{m+1}^+) - (SP_{m-1}^+ \cap SP_m^+) \\
 &\quad - (SP_m^+ \cap SP_{m+1}^+). \quad (54)
 \end{aligned}$$

C. Converting Overlapping Interval Into Shared Variables

Control variables at overlapping interval $T_{m-1} + 1$ are shared between SP_{m-1}^+ and SP_m^+ , and variables at interval $T_m + 1$ are shared between SP_m^+ and SP_{m+1}^+ . Each pair of shared variables must be the same to ensure consensus between subproblems and satisfy intertemporal constraints. We use two sets of control variables as shared variables: pre- and post-contingency variables. These sets change depending on the SCED model, whether it is preventive or corrective. Pre-contingency shared variables are power generated by thermal units, reserve up and down of thermal units, storage state of charge, and storage charge and discharge powers. Since corrective actions are considered in the formulated SCED, we model generators' power output and reserve, storage state of charge, and storage charge and discharge powers pertaining to each contingency as

post-contingency shared variables. The set of shared variables between SP_{m-1}^+ and SP_m^+ are:

$$\begin{aligned}
 \Xi_{m-1,m} &= \{p_{u,t}, r_{u,t}^{up}, r_{u,t}^d, E_{s,t}, p_{ch,s,t}, p_{dc,s,t}\} \\
 &\cup \{p_{u,t}^c, r_{u,t}^{up,c}, r_{u,t}^{d,c}, E_{s,t}^c, p_{ch,s,t}^c, p_{dc,s,t}^c\} \\
 &\quad \forall t = T_{m-1} + 1 \quad (55)
 \end{aligned}$$

Shared variables between SP_m^+ and SP_{m+1}^+ , $\Xi_{m,m+1}$, are the same as (55) but for $t = T_m + 1$.

D. Consistency Constraints

Shared variables between every two consecutive subproblems are duplicated. For instance, $\Xi_{m-1,m}$ is duplicated to create Ξ_{m-1} and Ξ_m . A set of consistency constraints is formulated and enforced in each SP to ensure that each pair of shared variables reaches the same value. The intertemporal consistency constraint between SP_{m-1}^+ and SP_m^+ are:

$$ICC_{m-1,m} : \Xi_{m-1} - \Xi_m = 0 \quad (56)$$

and they are as follows for SP_m^+ and SP_{m+1}^+ :

$$ICC_{m,m+1} : \Xi_m - \Xi_{m+1} = 0 \quad (57)$$

Ξ_m and $\Xi_{m'}$ denote the shared variables of subproblem m with its previous and next neighbors, respectively.

E. Storage Initial State of Charge

In the centralized SCED, the initial state of charge of storage is given. However, in the proposed temporal decomposition, the initial state of charge of storage is known only for the first SP, not others. This is an obstacle for formulating (16) in the first time interval of subproblems. One possible solution is to solve the first SP, pass the state of charge of storage obtained in its last time interval to the next SP, and use it as the initial state of charge of energy storage. This causes two problems: 1) the subproblems cannot be solved in parallel that increases the solution time, and 2) each SP forces its willingness for the state of charge of storage on its next subproblem that makes the resulted solution suboptimal. To solve this challenge and allow a parallel solution of subproblems while satisfying the solution optimality and feasibility, we introduce a new shared variable between every two consecutive subproblems. Consider SP_{m-1}^+ and SP_m^+ . The initial state of charge of storage in SP_m^+ needed for the overlapping interval $T_{m-1} + 1$ is the state of charge of storage at interval T_{m-1} . We introduce a control variable $E_{s,0}$ to allow SP_m^+ formulating (16) at its first interval $T_{m-1} + 1$. $E_{s,0}$ and its corresponding copy in SP_{m-1}^+ , i.e., $E_{s,T_{m-1}}$, are shared variables between these two subproblems. $E_{s,T_{m-1}}$ and $E_{s,0}$ are added to Ξ_{m-1} and Ξ_m , respectively, and the following intertemporal consistency constraint is included in both SCED subproblems:

$$E_{s,T_{m-1}} - E_{s,0} = 0. \quad (58)$$

IV. COORDINATION STRATEGY

Although enforcing (56)-(58) satisfies consistency between subproblems, it is an obstacle for solving subproblems in

parallel and makes the solution suboptimal. We relax these constraints in subproblems' objective functions using augmented Lagrangian relaxation and apply auxiliary problem principle (APP) to make mismatches between shared variables close to zero based on the concept of penalizing violations of consistency constraints. APP is a message passing-based iterative method that is suitable for parallel computing [11].

A. Initialization Strategy

APP, like most distributed optimization algorithms, is sensitive to the choice of initial values. We suggest an initialization strategy to enhance the performance of the coordination algorithm. The intertemporal coupling between subproblems is neglected, the overlapping time intervals are ignored, and subproblems are solved independently. The obtained values for variables corresponding to the overlapping time intervals are used for initialization. This enhances the performance of the coordination algorithm significantly as the obtained values for shared variables are usually near their optimal values, even if correlations between subproblems are ignored.

B. APP Implementation

The SCED subproblem SP_m^+ at iteration k is modified as follows:

$$\begin{aligned} \min \quad & \sum_t \sum_u a_u p_{u,t}^2 + b_u p_{u,t} + o_{u,t} + b_{ru} r_{u,t}^{up} \\ & + \sum_i b_{sh,i,t} p_{sh,i,t} + \frac{\rho}{2} \left\| \Xi_m^k - \Xi_m^{*k-1} \right\|^2 \\ & + \gamma \Xi_m^{k\dagger} \left(\Xi_m^{*k-1} - \Xi_m^{*k-1} \right) + \lambda^{(k-1)\dagger} \Xi_m^k \\ & + \frac{\rho}{2} \left\| \Xi_{m'}^k - \Xi_{m'}^{*k-1} \right\|^2 + \gamma \Xi_{m'}^{k\dagger} \left(\Xi_{m'}^{*k-1} - \Xi_{m+1}^{*k-1} \right) \\ & + \lambda'^{(k-1)\dagger} \Xi_{m'}^k \\ \text{s.t.} \quad & (3) - (38), (48) \& (49) \\ & \forall t = \{T_{m-1} + 1, \dots, T_m + 1\} \end{aligned} \quad (59)$$

where $\dagger$ is the transpose operator. The asterisk (*) refers to a variable whose value is determined by one of the neighboring subproblems and is kept fixed in subproblem m . SP_{m-1}^+ and SP_{m+1}^+ are formulated analogously with different time subhorizons. Another difference is in the last term of the penalty function that must be $-\lambda^{(k-1)\dagger} \Xi_m^k$ for SP_{m-1}^+ and $-\lambda'^{(k-1)\dagger} \Xi_{m'}^k$ for SP_{m+1}^+ . After each iteration, multipliers λ and λ' are updated as:

$$\lambda^k = \lambda^{k-1} + \omega \left(\Xi_{m-1}^{*k} - \Xi_m^{*k} \right) \quad (60)$$

$$\lambda'^k = \lambda'^{k-1} + \omega \left(\Xi_{m+1}^{*k} - \Xi_{m'}^{*k} \right) \quad (61)$$

where ω is a suitable constant (step size). Note that the value of the Lagrange multiplier λ in each iteration corresponds to the cost of maintaining the consistency constraints.

C. Discussion on Convergence

APP is proven to converge if subproblems are convex and the global optimal solution of each subproblem at each

iteration is obtained [11]. The objective function and constraints of the considered SCED problem are convex, except for (17)-(19) and (35)-(37) in which four sets of binary variables are defined for modeling storage charge/discharge status. These constraints do not add much complexity to the model. Solvers, such as CPLEX and Gurobi, are well advanced and provide a very high-quality solution for each subproblem. Therefore, at each iteration of APP, almost the global solution of each subproblem is obtained by the solver. Thus, after each iteration of APP, differences between shared variables decrease and the algorithm converges to the optimal point of the whole problem. It is also possible to convexify storage constraints by adding two small positive cost terms for storage charging and discharging (e.g., storage operation & maintenance costs) in subproblems objective functions and dropping binary variables [34]. This technique prevents simultaneous storage charge and discharge while making subproblems convex and ensuring APP convergence.

D. Guidelines on Number of Subproblems

Increasing the number of subhorizons for a given scheduling horizon results in smaller subproblems that are potentially less computationally expensive. This reduces the solution time of each iteration. However, increasing the number of subhorizons over a certain limit increases the required number of iterations for the distribution algorithm to converge. Hence, increasing the number of subhorizons is not necessarily efficient for time saving.

Our observations show that the load pattern has a significant impact on the optimal number of subhorizons. Decomposing the considered horizon from intervals with a low rate of change of load as compared to their neighboring intervals reduces the required number of iterations by the coordination algorithm to converge. Since the pattern of load is predictable in periodic, if an operator finds a good number of subhorizons for a load pattern, this information can be used for similar load patterns. Another factor that should be considered to take advantage of parallel computing is to make subproblems with similar size and computational complexity. This reduces the idle time of computing processors if synchronous coordination strategy, such as APP, is used. The strength of computing processors is another factor that should be considered. Making subhorizons smaller beyond a level may not lead to significant time saving as a computing processor could be strong enough to solve problems with X or 2X sizes within roughly the same time.

In a nutshell, we suggest the following steps to decompose a considered scheduling horizon. 1) Estimate processors computation time by reducing the size of the SCED problem. 2) Determine the size (assume it is X) beyond which the time reduction is not significant (steps 1 and 2 should be performed once, not once per load pattern). 3) Divide the size (assume it is NX) of the SCED problem into X to determine the number subhorizons (N). 4) Determine the lowest rate of change of load between consecutive intervals. 5) Decompose the scheduling horizon into N subhorizon from intervals with the lowest rate of change of load so that the size of subhorizons is similar.

V. CASE STUDY

The proposed algorithm is applied to solve a week-ahead SCED problem for a six-bus system, the IEEE 24-bus system, and a 472-bus system. System and equipment data are given in [35]. Simulations are carried out on MATLAB using YALMIP [36] as modeling software and Gurobi on a 3.7 GHz PC with 16GB of RAM. We have used one computing processor and have solved subproblems on this processor sequentially. To mimic parallel computing, when all subproblems are solved at each iteration of APP, the longest solution time is assumed as the runtime of that iteration. Upon convergence of the distributed algorithm, the runtimes of all iterations are summed up to determine the overall solution time.

A. Six-Bus Test System

This case serves as a tutorial for the proposed algorithm. The system includes six buses, three generating units, seven lines, three load points, one storage device, and a wind farm.

Validation of Nonparametric Chance Constraints for Reserve Procurement: The confidence level α is set to 0.95 for all intervals. To study the effectiveness of the proposed data-driven nonparametric chance constraints modeling with a reference point, we consider that the wind power follows a Gaussian PDF and generate 100 samples for each interval. The wind generation mean values over the scheduling horizon are given in [37], and the standard deviation is assumed to be 20%. Assuming that we do not know that wind generation follows Gaussian distribution, AKDE is applied to find nonparametric wind generation PDFs from the dataset. The pointwise error is calculated, and the adjusted risk levels of chance constraints are computed. For time interval 15, for instance, the squared pointwise error is 0.01 and according to (43), $d = 1e-04$. Plugging d in (44) yields $\alpha'^+_{15} = 0.048$. The resulted $CDF^{-1}_{\tilde{P}_{w,t}}(1 - \alpha'^+)$ and $CDF^{-1}_{\tilde{P}_{w,t}}(\alpha'^+)$ are 24.3376 and 23.5679, respectively. These values are used to formulate chance constraints and the centralized SCED is solved. The operation cost is \$391,225. As we assumed Gaussian PDFs for wind power, we can formulate classical parametric chance constraints and obtain the benchmark operation costs if complete information of wind power distributions is known. Table I shows that the relative error between the obtained costs is $3e-05$.

We have also assumed that wind generation follows Gamma distribution. The same procedure as of that for Gaussian distribution is implemented, and the results are depicted in Table II. The relative error is $3e-5$. The small relative errors reported in Tables I and II show that the proposed nonparametric approach can be adopted to formulate a data-driven SCED problem if only historical data is available and no information is known about the type of wind generation PDFs.

Sensitivity to Confidence Level: We have used three different values for α to study the performance of nonparametric models under different risk levels. It is assumed that wind power follows Gaussian distribution and 100 samples are generated for each interval. Table IV shows operation costs obtained by parametric and nonparametric models. For all three cases, the

TABLE I
COMPARISON BETWEEN PARAMETRIC AND NONPARAMETRIC CHANCE CONSTRAINTS USING GAUSSIAN DISTRIBUTION AS BENCHMARK

Method	SCED cost (\$)	Relative error
Parametric (benchmark)	\$391,213	-
Nonparametric	\$391,225	$3e-5$

TABLE II
COMPARISON BETWEEN PARAMETRIC AND NONPARAMETRIC CHANCE CONSTRAINTS USING GAMMA DISTRIBUTION

Method	SCED cost (\$)	Relative error
Parametric (benchmark)	399,054	-
Nonparametric	399,068	$3e-5$

TABLE III
OPERATION COSTS UNDER DIFFERENT RISK LEVELS

α	Parametric (benchmark)	Nonparametric	Relative error
0.1	\$39,1116	\$391,124	$2e-5$
0.05	\$391,213	\$391,225	$3e-5$
0.01	\$391,496	\$391,514	$5e-5$

relative error between the nonparametric model and benchmark results is acceptable, which shows that the proposed nonparametric chance-constrained programming works well under different risk levels.

Distributed SCED Analysis: The scheduling horizon is decomposed into seven equal-sized subhorizons, each of which includes 24 intervals. Line outage contingencies are considered. Adding the overlapping intervals results in a SCED subproblem with 25 time intervals for each subhorizon. There are 26 shared variables between every two consecutive SCED subproblems. The suggested initialization strategy is applied and its results are used to initialize APP. The distributed algorithm converges after ten iterations plus one initialization step. Fig. 4 shows the differences between shared variables over the course of iterations. Mismatches go to zero as more iterations are carried out. The majority of pairs of shared variables reach the consensus after six iterations. However, mismatches of Ξ_6 and Ξ_7 become less than the stopping threshold after ten iterations. As an example, Fig. 5 shows $p_{dc,1,25}^c$, storage discharge power after contingency one, from the perspective of SP_1 and $p_{dc,1,1}^c$ obtained by SP_2 . These two variables reach the same value after ten iterations.

Comparison With Centralized SCED: We have compared the values of decision variables obtained by the distributed and centralized SCED approaches. For instance, Figs. 6 (a) and (b) show generated powers by units one and two in the first subhorizon, and Figs. 6 (c) and (d) depict the storage discharge power and energy level in subhorizon seven. The optimal values obtained by the two approaches are almost the same. In addition, a convergence index is defined to measure the relative error between the operation costs determined by the distributed SCED (f^d) and the centralized SCED (f^{cp}), which is considered as benchmark results. The closer the convergence measure becomes to zero, the more precise solution

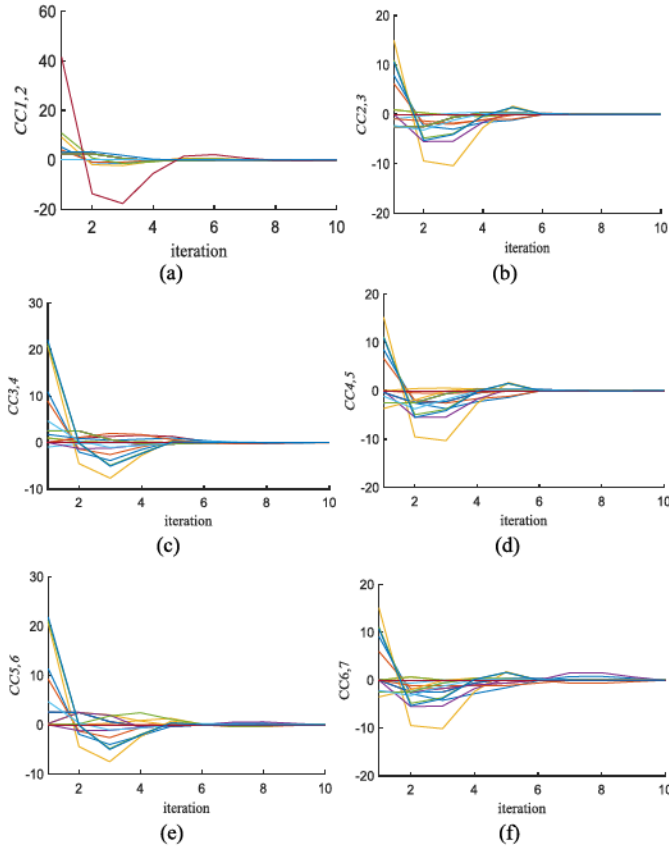


Fig. 4. Six-bus system: mismatches of consistency constraints between (a) SP_1 and SP_2 , (b) SP_2 and SP_3 , (c) SP_3 and SP_4 , (d) SP_4 and SP_5 , (e) SP_5 and SP_6 , and (f) SP_6 and SP_7 .

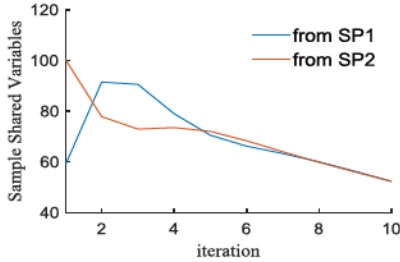


Fig. 5. Storage discharge power (MW) after contingency one.

is obtained.

$$rel = \frac{|f^{cp} - f^d|}{f^{cp}} \quad (62)$$

The convergence measure rel is depicted in Fig. 7. This index is $9e-05$ upon convergence, which shows the accuracy of the proposed distributed SCED algorithm.

Storage Initial State of Charge Modeling: Ignoring the initial state of charge consistency constraint (58), the distributed SCED is solved. The obtained solution is infeasible. For instance, the state of charge of storage at the last interval of SP_3 , determined by SP_3 , is 19 MWh. However, the initial state of charge of storage at the first interval of SP_4 , determined by SP_4 , is 100 MWh. Since these two values refer to the same physical variables, the obtained solution is

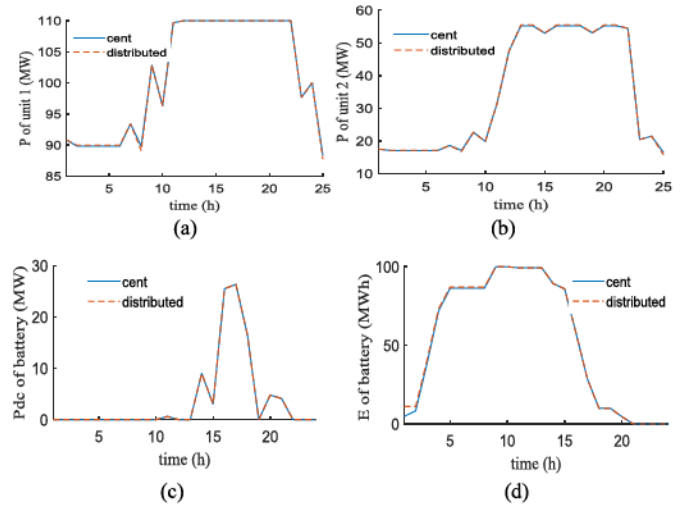


Fig. 6. The optimal values of (a) $p_{1,t}$, (b) $p_{2,t}$ in subhorizon one, and (c) $p_{dc,t}$, and (d) E_t in subhorizon seven.

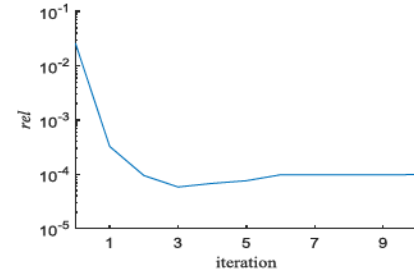


Fig. 7. The relative error for the six-bus system.

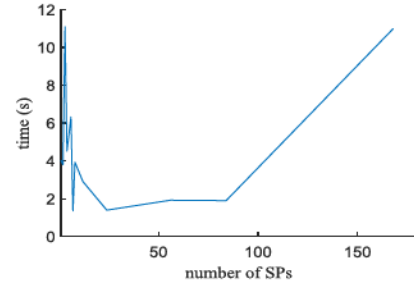


Fig. 8. Overall solution time versus number of subproblems.

infeasible. Incorporating (58) makes the solution feasible from the perspective of the whole scheduling horizon.

B. IEEE 24-Bus System

We have studied the impact of increasing the number of subhorizons on the solution time using the IEEE 24-bus system. We have increased the number of subproblems up to 168. Fig. 8 shows the solution time versus the number of subproblems. Increasing the number of subproblems up to 24 results in the overall solution time reduction (although some jumps are observed, the general pattern is decreasing). While increasing the number of subproblems from 24 to 84 does not have a considerable impact on the solution time, increasing the number of subproblems beyond 84 increases the solution time. Such a curve with this trend is reported in parallel computing literature.

TABLE IV
OPERATION COSTS UNDER DIFFERENT APP MULTIPLIER ω

Algorithm	ω	Cost (\$)	rel	Time (s)	Time saving
Centralized	-	41,281,398	-	654	-
	1	41,283,289	4e-5	312	53%
Distributed	0.2	41,281,975	1e-5	346	47%
	0.05	41,281,401	2e-8	396	40%

C. 472-Bus Test System

This system, which is created by connecting four IEEE 118-bus systems, has 216 generators, 784 lines, 364 load points, one storage device, and a wind farm. The confidence levels of chance constraints are set to 0.95. The size of the centralized problem is relatively large. As illustrated in Table IV, the solver takes 654 seconds to solve the centralized SCED and return the optimal operation cost of \$41,281,398.

We partition the overall time horizon into seven subhorizons and implement the proposed algorithm. We have selected three values for step size ω for updating Lagrange multipliers. Increasing the value of ω speeds up the solution procedure. However, it affects the relative error. In general, the larger ω is, the faster the algorithm would converge and the larger the rel index would be. While $\omega = 1$ yields the least solution time, 312 seconds, $\omega = 0.05$ results in the least rel index, almost zero. Comparing the solution time and rel , we select $\omega = 0.05$. The distributed algorithm converges after 11 iterations within 396 seconds. The proposed distributed SCED algorithm is 40% faster than the centralized SCED, while the operation costs obtained by both algorithms are almost the same.

D. Hybrid Geographical and Temporal Decomposition

We demonstrate that temporal decomposition can be combined with geographical decomposition to enhance the performance of distributed SCED. A 944-bus test system, whose information is given in [35], is used. The considered scheduling horizon includes 168 intervals. For the sake of explanation and simplicity, we have ignored energy storage, wind generation, load shedding, and contingency. As shown in Table V, the centralized approach takes 46 seconds to solve the problem, and the total cost is \$88,749,812. The system is decomposed into two regions, and geographical decomposition is applied [38]. The distributed SCED converges after 31 seconds with $rel \approx 0$. The proposed temporal decomposition is applied. Each subhorizon includes 24 intervals. The algorithm takes 25 seconds to converge to $rel \approx 0$.

We have combined geographical and temporal decomposition strategies to solve SCED in a distributed manner. This hybrid approach has two loops, an inner loop and an outer loop. In the outer loop, subproblems obtained by geographical decomposition are solved, and the corresponding Lagrange multipliers are updated iteratively using APP. Whereas in the inner loop, each subproblem of geographical decomposition is further decomposed over the time horizon, and APP is applied to coordinate subproblems. To have a fair comparison, similar to solely geographical decomposition and solely temporal decomposition, we have considered two zones and

TABLE V
COMPARISON BETWEEN DIFFERENT APPROACHES TO SOLVE SCED FOR THE 994-BUS SYSTEM

Method	Time (sec)	rel
Centralized (benchmark)	46	-
Geographical decomposition	31	≈ 0
Temporal decomposition	25	≈ 0
Hybrid geographical + Temporal	10	≈ 0

seven subhorizons. The hybrid decomposition strategy outperforms both geographical and temporal decompositions and converges within ten seconds with $rel \approx 0$. This shows that the temporal decomposition can be combined with geographical decomposition to enhance the performance of the distributed SCED.

VI. CONCLUSION

A temporal decomposition approach is proposed to decompose the SCED problem over the scheduling horizon with the goal of reducing the computational burden of the optimization problem. Two data-driven nonparametric chance constraints are formulated for reserve up and down requirements considering wind generation uncertainty. These constraints are replaced by their equivalent parametric constraints to make SCED solvable by standard solvers. Generators' ramp limits and the state of charge of storage for the transition between subhorizons are modeled by introducing overlapping intervals and 12 sets of shared variables. It is discussed that the initial state of charge of storage at the beginning of each subproblem must be the same as the state of charge of storage at the interval before the overlapping interval; otherwise, the solution will be infeasible. An initialization strategy is suggested to enhance the performance of the distributed coordination algorithm.

A tutorial is presented based on a small system, the impact of number of subhorizons on the algorithm is studied using the IEEE 24-bus system, and the results for a 472-bus system show that the SCED solution time is reduced by a factor of 1.65 as compared to that of the centralized SCED. As the size of the optimization problem increases, the effectiveness of the proposed method is more considerable. In addition, it is illustrated that data-driven nonparametric chance constraints provide a solution close to the benchmark results obtained using the complete information of wind generation density function.

A direction of research is to combine reserve up (or down) requirements at all intervals in a probabilistic constraint and develop a nonparametric joint chance constraint model. Another possible research direction is to develop methods to find an optimal time decomposition strategy and to improve the convergence performance of the coordination algorithm by using momentum and second-order derivative information for Lagrange multipliers updating.

REFERENCES

- [1] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control*. Hoboken, NJ, USA: Wiley, 2013.

- [2] A. Kargarian *et al.*, "Toward distributed/decentralized DC optimal power flow implementation in future electric power systems," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2574–2594, Jul. 2018.
- [3] C. Lin, W. Wu, X. Chen, and W. Zheng, "Decentralized dynamic economic dispatch for integrated transmission and active distribution networks using multi-parametric programming," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4983–4993, Sep. 2018.
- [4] M. H. Amini, R. Jaddivada, S. Mishra, and O. Karabasoglu, "Distributed security constrained economic dispatch," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT ASIA)*, 2015, pp. 1–6.
- [5] J. Hu, M. Z. Q. Chen, J. Cao, and J. M. Guerrero, "Coordinated active power dispatch for a microgrid via distributed lambda iteration," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 2, pp. 250–261, Jun. 2017.
- [6] D. K. Molzahn *et al.*, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2941–2962, Nov. 2017.
- [7] Y. Wang, S. Wang, and L. Wu, "Distributed optimization approaches for emerging power systems operation: A review," *Elect. Power Syst. Res.*, vol. 144, pp. 127–135, Mar. 2017.
- [8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [9] H. M. Kim, N. F. Michelena, P. Y. Papalambros, and T. Jiang, "Target cascading in optimal system design," *J. Mech. Design*, vol. 125, no. 3, pp. 474–480, 2003.
- [10] A. J. Conejo, F. J. Nogales, and F. J. Prieto, "A decomposition procedure based on approximate Newton directions," *Math. Program.*, vol. 93, no. 3, pp. 495–515, 2002.
- [11] G. Cohen, "Auxiliary problem principle and decomposition of optimization problems," *J. Optim. Theory Appl.*, vol. 32, no. 3, pp. 277–305, 1980.
- [12] H. Xing, Y. Mou, M. Fu, and Z. Lin, "Distributed bisection method for economic power dispatch in smart grid," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3024–3035, Nov. 2015.
- [13] V. Loia and A. Vaccaro, "Decentralized economic dispatch in smart grids by self-organizing dynamic agents," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 4, pp. 397–408, Apr. 2014.
- [14] Y. Zhang and G. B. Giannakis, "Efficient decentralized economic dispatch for microgrids with wind power integration," in *Proc. 6th Annu. IEEE Green Technol. Conf.*, 2014, pp. 7–12.
- [15] Q. Li, D. W. Gao, H. Zhang, Z. Wu, and F.-Y. Wang, "Consensus-based distributed economic dispatch control method in power systems," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 941–954, Jan. 2019.
- [16] W. T. Elsayed and E. F. El-Saadany, "A fully decentralized approach for solving the economic dispatch problem," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2179–2189, Jul. 2015.
- [17] F. Guo, C. Wen, J. Mao, and Y.-D. Song, "Distributed economic dispatch for smart grids with random wind power," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1572–1583, May 2016.
- [18] S. Yang, S. Tan, and J.-X. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.
- [19] I. U. Nutkani, P. C. Loh, P. Wang, and F. Blaabjerg, "Decentralized economic dispatch scheme with online power reserve for microgrids," *IEEE Trans. Smart Grid*, vol. 8, no. 1, pp. 139–148, Jan. 2017.
- [20] Y. Zhang, Y. Sun, X. Wu, D. Sidorov, and D. Panasetzky, "Economic dispatch in smart grid based on fully distributed consensus algorithm with time delay," in *Proc. 37th Chin. Control Conf. (CCC)*, 2018, pp. 2442–2446.
- [21] Y. Gu and L. Xie, "Stochastic look-ahead economic dispatch with variable generation resources," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 17–29, Jan. 2017.
- [22] W. Wu, J. Chen, B. Zhang, and H. Sun, "A robust wind power optimization method for look-ahead power dispatch," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 507–515, Apr. 2014.
- [23] R. Khatami, M. Parvania, and P. P. Khargonekar, "Continuous-time look-ahead scheduling of energy storage in regulation markets," in *Proc. 52nd Hawaii Int. Conf. Syst. Sci.*, 2019, pp. 1–9.
- [24] A. Ahmadi-Khatir, A. J. Conejo, and R. Cherkaoui, "Multi-area unit scheduling and reserve allocation under wind power uncertainty," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1701–1710, Jul. 2014.
- [25] A. Kargarian, Y. Fu, and H. Wu, "Chance-constrained system of systems based operation of power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3404–3413, Sep. 2016.
- [26] J. Yu, Z. Li, Y. Guo, and H. Sun, "Decentralized chance-constrained economic dispatch for integrated transmission-district energy systems," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6724–6734, Nov. 2019.
- [27] M. B. Cain, R. P. O'Neill, and A. Castillo, "History of optimal power flow and formulations," *Federal Energy Regulatory Commission*, vol. 1, pp. 1–36, Dec. 2012.
- [28] Y. Wen, C. Guo, H. Pandžić, and D. S. Kirschen, "Enhanced security-constrained unit commitment with emerging utility-scale energy storage," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 652–662, Jan. 2016.
- [29] H. Wu, M. Shahidehpour, Z. Li, and W. Tian, "Chance-constrained day-ahead scheduling in stochastic power system operation," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1583–1591, Jul. 2014.
- [30] B. A. Calfa, I. E. Grossmann, A. Agarwal, S. J. Bury, and J. M. Wassick, "Data-driven individual and joint chance-constrained optimization via kernel smoothing," *Comput. Chem. Eng.*, vol. 78, pp. 51–69, Jul. 2015.
- [31] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Math. Program.*, vol. 158, nos. 1–2, pp. 291–327, 2016.
- [32] L. Pardo, *Statistical Inference Based on Divergence Measures*. Boca Raton, FL, USA: CRC Press, 2005.
- [33] P. V. Kerm, "Adaptive kernel density estimation," *Stata J.*, vol. 3, no. 2, pp. 148–156, 2003.
- [34] P. Yang and A. Nehorai, "Joint optimization of hybrid energy storage and generation capacity with renewable energy," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1566–1574, Jul. 2014.
- [35] Accessed: Apr. 7, 2020. [Online]. Available: <https://sites.google.com/site/aminkargarian/test-system-data/temporal-decomposition-based-stochastic-economic-dispatch-for-smart-grid>
- [36] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. Comput. Aided Control Syst. Design*, 2004, pp. 284–289.
- [37] J. Guo, G. Hug, and O. K. Tonguz, "Intelligent partitioning in distributed optimization of electric power systems," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1249–1258, May 2016.
- [38] M. Mehrtash, A. Kargarian, and A. Mohammadi, "Distributed optimisation-based collaborative security-constrained transmission expansion planning for multi-regional systems," *IET Gener. Transm. Distrib.*, vol. 13, no. 13, pp. 2819–2827, 2019.

Farnaz Safdarian (Student Member, IEEE) received the B.Sc. degree in electrical engineering (Power) from Shahid Beheshti University, Tehran, Iran, in 2011, and the M.Sc. degree (First Rank Hons.) in electrical engineering (Power) from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2014. She is currently pursuing the Ph.D. degree in electrical engineering (Power) with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA. Her research interests include distributed optimization, power systems operation and planning, machine learning and data science, and electricity markets.

Amin Kargarian (Member, IEEE) received the B.Sc. degree in electrical and computer engineering from the University of Isfahan, Isfahan, Iran, in 2007, the M.Sc. degree in electrical and computer engineering from Shiraz University, Shiraz, Iran, in 2010, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2014. In 2014, he joined the ECE Department, Carnegie Mellon University as a Postdoctoral Research Associate. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA, USA. His research interests include power system optimization and machine learning.