

# A Fast Solution Method for Stochastic Transmission Expansion Planning

Junpeng Zhan, *Member, IEEE*, C. Y. Chung, *Fellow, IEEE*, and Alireza Zare, *Student Member, IEEE*

**Abstract**—Stochastic programming is a cost-effective approach to model the transmission expansion planning (TEP) considering the uncertainties of wind and load, which is known as stochastic TEP (STEP). The uncertainty can be accurately represented by a large number of scenarios which need to be reduced to a relatively small number in order to shorten the computational time required by the STEP. The forward selection algorithm (FSA) is an accurate scenario reduction method which, however, is quite time consuming. An improved FSA (IFSA) is proposed in order to shorten the computational time. The STEP is a large-scale mixed-integer programming problem and therefore is difficult to be solved directly. Benders decomposition algorithm is suitable to solve the STEP by decomposing it into master and multiple slave problems. The slave problems are nonlinear and thereby are difficult and time consuming to be solved. In this regard, a linearization method is proposed to solve the slave problems faster and to calculate the Lagrangian multipliers needed by the master problem. Two medium and a large data sets are used to demonstrate the efficiency of the IFSA and a 24-, a 300-, and a 2383-bus test systems are used to verify the efficiency of the linearization method.

**Index Terms**—Benders decomposition algorithm; Lagrangian multipliers; linearization; scenario reduction method; stochastic programming; transmission expansion planning (TEP).

## I. NOMENCLATURE

The terminologies used in this paper are listed as follows:

### Sets/Indices

$l$	Line index
$i$	Bus index
$n$	Candidate line index
$g$	Generation
$r$	Load loss
$s$	Wind spillage
$\kappa, \omega, x$	Scenario
$\xi$	Iteration in Benders decomposition
$\Omega$	Set of all existing and candidate lines
$\Omega_0$	Set of all existing lines
$\Omega_1$	Set of $l$ where there is at least one line in operation, no matter initially or newly installed, on right-of-way $l$
$\Omega_{os}$	Original set of scenarios
$\Omega_{rs}$	Reduced set of scenarios
$\Omega_B$	Set of all buses

$\Omega_J$	Original set of scenarios excluding the reduced set of scenarios
$\Omega_N$	Set of all candidate lines
$\mathcal{C}$	Cluster centers obtained by the K-means
$\mathcal{C}_j$	All the elements in the $j$ th cluster obtained by the K-means

### Parameters

$c_l$	Cost of a line added to right-of-way $l$ (\$)
$d_i^k$	Electricity load demand at bus $i$ (MW)
$f_l^k$	Total active power flow on right-of-way $l$ (MW)
$f_l^{max}$ , $f_n^{max}$	The maximum power flow of line $l$ and $n$ , respectively (MW)
$\bar{g}_i$	Maximum power output of generator $i$ (MW)
$h$	Number of hours in the study period (hour)
$k$	Number of clusters in the K-means and the number of initial clusters in the first step of the improved forward selection algorithm
$\bar{n}_l$	Maximum number of new lines can be added to right-of-way $l$ (set to 3 in this paper)
$n_l^0$	Initial number of lines on right-of-way $l$
$n_{rs}$	Number of reduced scenarios
$p^\kappa, p^\omega$ , $p^x$	Probability of scenarios $\kappa$ , $\omega$ , and $x$ , respectively
$x^c, x_i^c$	A cluster center in the K-means, also as a scenario in $\mathcal{C}$ , $i = 1, 2, \dots, k$
$B_M$	Large value for a disjunctive constraint
$L$	Total number of right-of-ways
$M_{i,l}$	The element in the $i$ th row and the $l$ th column of node-branch incidence matrix
$N_{os}$	Total number of elements in the original set $\Omega_{os}$
$N_{rs}$	Total number of elements in the reduced set $\Omega_{rs}$
$\bar{W}_i^k$	Maximum wind power that can be generated at bus $i$ (MW)
$Z_{up}^\xi$	Upper bound obtained after solving all the slave problems in iteration $\xi$
$Z_{down}^\xi$	Lower bound obtained after solving the master problem in iteration $\xi$
$\tilde{\gamma}_l$	Susceptance of a line on right-of-way $l$ (Siemens)
$\theta_{l,f}^k$	Phase angle of from-side node of right-of-way $l$ (rad)
$\theta_{l,t}^k$	Phase angle of to-side node of right-of-way $l$ (rad)
$\kappa$	The superscript represents scenario $\kappa$ associated with wind and electricity load demand in different time of a year
$\lambda_{1,i}^\kappa, \lambda_{1,i}^{\kappa'}$ , $\lambda_{2,l}^\kappa$	Lagrangian multipliers associated with equality constraints (15b), (19), and (16b), respectively

The work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Saskatchewan Power Corporation (SaskPower).

Junpeng Zhan, C. Y. Chung, and Alireza Zare are with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: [j.p.zhan@usask.ca](mailto:j.p.zhan@usask.ca), [c.y.chung@usask.ca](mailto:c.y.chung@usask.ca), [alireza.zare@usask.ca](mailto:alireza.zare@usask.ca)).

$\mu_{i,l}^{\kappa}$	Lagrangian multipliers associated with inequality constraints (15c)-(15j), $i = 1, 2, \dots, 8$
$\xi$	The superscript represents the iteration in the Benders decomposition algorithm
$\rho_g$	Vector of generation costs of generators (\$/MWh)
$\rho_r$	Penalty cost for electricity load loss (\$/MWh)
$\rho_s$	Penalty cost for wind spillage (\$/MWh)
$\varphi^{\kappa}$	Objective function of the $\kappa$ th slave problem
$\varphi^{\kappa,\xi}$	Objective function of the $\kappa$ th slave problem in iteration $\xi$

**Variables**

$g_i^{\kappa}, \mathbf{g}^{\kappa}$	Power output of generator at bus $i$ , and the corresponding vector (MW)
$n_l$	Number of new lines added to right-of-way $l$
$r_i^{\kappa}, \mathbf{r}^{\kappa}$	Electricity load loss at bus $i$ , and the corresponding vector (MW)
$x_n$	Binary decision variable indicating whether or not to add a new line
$S_i^{\kappa}, \mathbf{S}^{\kappa}$	Wind power curtailment at bus $i$ , and the corresponding vector (MW)
$\beta$	A decision variable in a master problem

## II. INTRODUCTION

TRANSMISSION expansion planning (TEP) problem is to ensure sufficient transmission capacity for the future load of a power system at a minimum cost. Traditionally, TEP targets at satisfying the maximum load of a future year, which has been researched for decades. A comprehensive review about TEP in both regulated and deregulated environments has been given in [1], [2].

Recently, a large amount of renewables has been integrated into power systems. The wind power is one of the most important renewables, which is growing very fast. According to Global Wind Energy Council [3], the installed wind power could reach 2,000 GW by 2030 (about 1.8 times as much as the value in 2015), and supply up to 17-19% of global electricity. By 2050, wind power could provide 25-30% of global electricity supply. That is, a growing share of electricity will be supplied by wind power in the future. Consequently, it is necessary to accommodate the wind power in TEP [4]-[6], which is important but difficult due to its intermittency and unpredictability.

Recently, several new methods have been proposed to solve the TEP problem considering the uncertainties of wind and load, e.g., a stochastic programming [4], [5], a robust optimization [7], a chance-constrained method [8], a probabilistic branch and bound method [9], etc. According to the analysis given in [10], the stochastic programming is a more cost-effective way to deal with the wind uncertainty in comparison with the robust and the interval optimizations. Thus, the stochastic programming is adopted in this paper for the TEP considering uncertainty, i.e., the TEP becomes a stochastic TEP (STEP) problem.

The uncertainty can be accurately represented by a large number of scenarios. However, using a large number of scenarios in stochastic programming is intractable as a very long computational time is required. Thus, it is necessary to reduce a large number of original scenarios into a relatively small number of scenarios [11]-[13].

Due to the importance of scenarios reduction, much work has focused on this area and it is still an active research area in order to reduce a large number of scenarios in an accurate and fast manner [14]. Existing scenario reduction methods include clustering algorithms such as K-means [15], backward and forward selection algorithms [11], [12], and their variation or extension [11], [16], heuristic-based methods [14], [17], etc. Among them, forward selection algorithm (FSA) [11]-[13] is one of the most widely used scenario reduction methods, which can generate reduced scenarios that perform well in practice [13]. However, the FSA is very time consuming when used to reduce a large number of scenarios.

In order to address this issue, an improved FSA (IFSA) is proposed in this paper. In the FSA, in order to select a scenario to be added to the reduced set of scenarios, each element in the original set of scenarios needs to be traversed, which is quite time consuming. The idea of the IFSA is to divide the original set of scenarios into a number of clusters and thereafter all the cluster centers are traversed to select the best center. Then, each element in the single cluster associated with the best center is traversed. Besides, a strategy is proposed in the IFSA such that the calculation result obtained in the previous iteration can be utilized in the current iteration, which can avoid a large amount of redundant calculations. Therefore, the computational burden of IFSA is significantly less than that of FSA.

It is worth to discuss a stratified sampling method named Latin hypercube sampling (LHS) [18], which provides a simple and effective means of generating a limited number of scenarios without scenario reduction. In the LHS, the marginal cumulative distribution function of each random variable is divided into equal strata and then a value is randomly picked from each stratum. In the second step, the sampled values for each marginal distribution are permuted to create scenarios based on the correlation between the random variables. The second step of the LHS is crucial as it needs to generate correlated random variables which should have the same correlation characteristics as the original random variables. In reality, the correlation pattern between different random variables is usually not a standard copula, which increases the difficulty of the second step of the LHS. To address this issue, several new technologies [19]-[21] have been proposed recently. The advantage of scenario reduction methods such as the FSA over the LHS lies in that it does not need to know the correlation between different random variables.

The main mathematical formulations used in TEP, including a DC model, a transportation model, a hybrid model, and a disjunctive model, have been summarized and compared in [22]. The DC model is a mixed-integer and nonlinear model, which is difficult to be solved. Usually, the DC model is solved by evolutionary algorithms [8], [23], which cannot guarantee obtaining the optimal solution and may not be efficient in a large system. The optimal solution of the transportation model is not necessarily feasible for the DC model and consequently may have higher investment cost than the optimal solution of the DC model. The hybrid model does not consider the Kirchhoff's voltage law (KVL) constraint for the newly added circuits. The transportation model and the hybrid model are usually not used as they are inaccurate. The disjunctive model is a linear model and keeps the same accuracy as the DC model,

which has been adopted in [5], [24] to solve a stochastic transmission planning model considering the uncertainties of wind and load. The disadvantage of the disjunctive model lies in the fact that it uses many more variables and inequality constraints compared to the DC model [22].

In this paper, the DC model is adopted for STEP. The STEP is decomposed into a master problem and multiple slave problems by using the Benders decomposition algorithm, where the slave problems are nonlinear. A linearization method is proposed to solve the slave problems as linear programming models and to calculate the Lagrangian multipliers needed by the master problem. The advantage of this linearization method is that the DC model is converted to a linear model which can be solved significantly faster than the original non-linear DC model. Besides, the DC model has fewer variables and inequality constraints than the disjunctive model and consequently the former can be solved faster than the latter, especially in a large system.

The contribution of this paper is twofold. On the one hand, a new scenario reduction method, i.e., the IFSA, has been proposed, which consumes a significantly shorter computational time than the FSA but is almost as accurate as the FSA. On the other hand, a linearization method has been proposed to solve the slave problems in the Benders decomposition algorithm for the STEP problem, which is accurate and much faster than existing methods.

The rest of the paper is organized as follows. The IFSA is described in Section III. In Section IV, the STEP is decomposed into a master problem and multiple slave problems and then the proposed linearization method for solving the slave problems is detailed. Simulation results are given in Section V. Conclusions are given in Section VI. At last, a proof and a model are provided in Appendix.

### III. SCENARIOS REDUCTION

Uncertainty can be accurately represented by a large number of scenarios, which usually results in a very long computational time. To address this issue, scenario reduction is an important topic, which is to find a reduced set,  $\Omega_{rs}$ , that has the minimal probability distance to the original set,  $\Omega_{os}$ . The Kantorovich distance is the most common probability distance used in stochastic programming, which can be expressed as [13]:

$$D_K(\Omega_{os}, \Omega_{rs}) = \sum_{\omega \in \Omega_{os} \setminus \Omega_{rs}} p^\omega \min_{\omega' \in \Omega_{rs}} c(\omega, \omega'). \quad (1)$$

where  $D_K(\Omega_{os}, \Omega_{rs})$  is the Kantorovich distance between  $\Omega_{os}$  and  $\Omega_{rs}$ ;  $c(\omega, \omega')$  is the Euclidean distance between scenarios  $\omega$  and  $\omega'$ ;  $\Omega_{os} \setminus \Omega_{rs}$  represents the set of  $\Omega_{os}$  excluding  $\Omega_{rs}$ .

FSA is a widely used scenario reduction method, which can generate reduced sets that perform well in practice [13]. This method is an iterative greedy process to select a reduced set which has the minimal Kantorovich distance to the original set. The main drawback of the FSA is that it is very time consuming when the original set has a large number of elements. In the following, firstly, it is indicated that the K-means clustering method can be used for scenario reduction. Then, the IFSA is detailed.

#### A. Relationship Between the K-means and the FSA

In the K-means clustering method, we are given an integer  $k$  and a set of data points  $\chi$ . We wish to choose a set of  $k$  centers,  $\mathcal{C} = \{x_1^c, x_2^c, \dots, x_k^c\}$ , so as to minimize the potential function  $\Phi$  [25]:

$$\Phi = \sum_{x \in \chi} \min_{x^c \in \mathcal{C}} \|x - x^c\|^2, \quad (2)$$

where  $\|x - x^c\|^2$  is the Euclidean distance between  $x$  and  $x^c$ . Let  $\omega = x$ ,  $\Omega_{os} = \chi$ ,  $\omega' = x^c$ ,  $\Omega_{rs} = \mathcal{C}$ , and  $c(\omega, \omega') = \|\omega - \omega'\|^2$ . Then (2) can be written as:

$$\Phi = \sum_{\omega \in \Omega_{os}} \min_{\omega' \in \Omega_{rs}} c(\omega, \omega'). \quad (3)$$

In the K-means clustering method, the probability of each original scenario is the same, i.e.,  $p^\omega$  is the same for all  $\omega$ . Then, after multiplying each term in both sides of (3) by  $p^\omega$ , we can obtain

$$p^\omega \Phi = \sum_{\omega \in \Omega_{os}} p^\omega \min_{\omega' \in \Omega_{rs}} c(\omega, \omega'). \quad (4)$$

Note that

$$\sum_{\omega \in \Omega_{rs}} p^\omega \min_{\omega' \in \Omega_{rs}} c(\omega, \omega') = 0. \quad (5)$$

By adding (5) to (1), we can obtain:

$$D_K(\Omega_{os}, \Omega_{rs}) = \sum_{\omega \in \Omega_{os}} p^\omega \min_{\omega' \in \Omega_{rs}} c(\omega, \omega'). \quad (6)$$

Comparing (6) with (4), it can be known that  $D_K(\Omega_{os}, \Omega_{rs}) = p^\omega \times \Phi$ . That is, when each original scenario has the same probability, the objective of the FSA is equal to  $p^\omega$  times the objective of the K-means. Then, we can use the K-means to solve (2) to obtain  $\mathcal{C}$  which can be considered as the reduced set  $\Omega_{rs}$  in (1). The difference between  $\mathcal{C}$  and  $\Omega_{rs}$  lies in the fact that each element in  $\Omega_{rs}$  belongs to  $\Omega_{os}$  while the elements in  $\mathcal{C}$  may not belong to  $\Omega_{os}$ . For the purpose of comparison, the K-means clustering method is also used for scenario reduction in this paper.

#### B. Main Idea of the IFSA

According to [13], in step  $i$  of the FSA, which is the main step, a new scenario is added to the reduced scenario set. The selection of this scenario is carried out using

$$\omega_i = \arg \left\{ \min_{\omega' \in \Omega_j^{[i-1]}} \sum_{\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}} p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]} \cup \{\omega'\}} c(\omega, \omega'') \right\} \quad (7)$$

where  $\Omega_j^{[i]}$  represents the set consisting of those scenarios which have not been selected in the first  $i$  steps of the FSA and  $\Omega_{rs}^{[i]}$  represents the set consisting of the selected scenarios until step  $i$ . Note that  $\Omega_j^{[i]} \cup \Omega_{rs}^{[i]} = \Omega_{os}$ ,  $\Omega_{rs}^{[0]} = \emptyset$ ,  $\Omega_j^{[i]} = \Omega_j^{[i-1]} \setminus \{\omega_i\}$ , and  $\Omega_{rs}^{[i]} = \Omega_{rs}^{[i-1]} \cup \{\omega_i\}$ .

In the FSA, in order to obtain  $\omega_i$  in (7), each element in  $\Omega_j^{[i-1]}$  is traversed. When the number of elements in  $\Omega_j^{[i-1]}$  is large, step  $i$  is very time consuming. The idea of the IFSA is to divide the original set  $\Omega_{os}$  into a number of clusters, then traverse the center of each cluster to find the best center, and thereafter traverse all the elements in the cluster associated with the best center. That is, in the IFSA, all the cluster centers and all the elements in the single cluster associated with the best center are traversed, which requires significantly less

computational effort compared to traversing all the elements in all the clusters in the FSA.

In the IFSA, through traversing all the cluster centers of the original set of scenarios, the best center is selected as

$$\omega_{i,j} = \arg \left\{ \min_{\omega' \in \mathcal{C}} \sum_{\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}} p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]} \cup \{\omega'\}} c(\omega, \omega'') \right\} \quad (8)$$

where  $\omega_{i,j}$  represents the  $j$ th element in  $\mathcal{C}$ , i.e., the best center.

Then, all the elements in the  $j$ th cluster of the original set of scenarios are traversed:

$$\omega_i = \arg \left\{ \min_{\omega' \in \mathcal{C} \cup \mathcal{C}_j} \sum_{\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}} p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]} \cup \{\omega'\}} c(\omega, \omega'') \right\} \quad (9)$$

where  $\mathcal{C}_j$  represents all the elements in the  $j$ th cluster.

### C. Strategy to Avoid Redundant Calculations in the IFSA

In the following, an equivalent expression of (8) and (9) is deduced, which can avoid a large amount of redundant calculations and thereby can significantly reduce the computational burden required by (8) and (9).

Note that  $p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]} \cup \{\omega'\}} c(\omega, \omega'')$  is equivalent to  $\min \left\{ p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]}} c(\omega, \omega''), p^\omega \min_{\omega'' \in \{\omega'\}} c(\omega, \omega'') \right\}$ . Then, (8) and (9) can be respectively rewritten as

$$\omega_{i,j} = \arg \left\{ \min_{\omega' \in \mathcal{C}} \sum_{\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}} \min \left\{ p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]}} c(\omega, \omega''), p^\omega \min_{\omega'' \in \{\omega'\}} c(\omega, \omega'') \right\} \right\}, \quad (10)$$

$$\omega_i = \arg \left\{ \min_{\omega' \in \mathcal{C} \cup \mathcal{C}_j} \sum_{\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}} \min \left\{ p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]}} c(\omega, \omega''), p^\omega \min_{\omega'' \in \{\omega'\}} c(\omega, \omega'') \right\} \right\}. \quad (11)$$

Note that  $\omega \in \Omega_j^{[i-1]} \setminus \{\omega'\}$  in (10) and (11) can be replaced by  $\omega \in \Omega_j^{[0]} = \Omega_{os}$  and that  $p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]}} c(\omega, \omega'')$  is a value obtained in the  $(i-1)$ th iteration, which can be stored and used in the  $i$ th iteration. Therefore, in the  $i$ th iteration, only  $p^\omega \min_{\omega'' \in \{\omega'\}} c(\omega, \omega'')$  needs to be calculated. The difference between (10)-(11) and (8)-(9) is whether  $p^\omega \min_{\omega'' \in \Omega_{rs}^{[i-1]}} c(\omega, \omega'')$  is calculated in the  $i$ th iteration or not.

The ratio of the computational burden in the  $i$ th iteration associated with (8)-(9) to that associated with (10)-(11) is  $\left\| \Omega_{rs}^{[i-1]} \cup \{\omega'\} \right\| / \left\| \{\omega'\} \right\|$  which is equal to  $\left\| \Omega_{rs}^{[i-1]} \right\| + 1$ . Then, the ratio of the total computational burden associated with (8)-(9) to that associated with (10)-(11) is  $(1 + 2 + \dots + n_{rs}) / n_{rs}$  which is equal to  $(1 + n_{rs}) / 2$ . That is, by using (10)-(11) instead of (8)-(9), the computational burden is reduced by  $(1 + n_{rs}) / 2$  times.

### D. Procedure of the IFSA

The procedure of the IFSA is as follows:

**Step 1:** Divide the original set of scenarios into  $k$  clusters. The cluster centers are denoted as  $\mathcal{C} = \{x_1^c, x_2^c, \dots, x_k^c\}$ , where  $k$  is called the number of initial clusters for the convenience of

expression. In this step, a widely used clustering method, the K-means clustering method, is employed. Note that other clustering methods can also be used in this step.

**Step 2:** Select the starting scenario  $\omega_1$  by [13]:

$$\omega_1 = \arg \left\{ \min_{\omega' \in \Omega_{os}} \sum_{\omega \in \Omega_{os}} p^\omega c(\omega, \omega') \right\} \quad (12)$$

which is the same as the first step of the FSA.

**Step 3:** Select an element from the original set of scenarios to add to the reduced set of scenarios by using (11), where  $j$  is determined in (10).

**Step 4:** Repeat step 3 until enough elements are added to the reduced set of scenarios. The obtained reduced set is denoted as  $\Omega_{rs}^*$ .

**Step 5:** Calculate the probability of each element in the reduced set via

$$p^{x^*} \leftarrow \sum_{x \in J(x^*)} p^x, \quad \forall x^* \in \Omega_{rs}^*, \quad (13)$$

where  $J(x^*) = \left\{ x \in \Omega_{os} \mid x^* \in \arg \min_{x'' \in \Omega_{rs}^*} c(x'', x) \right\}$ . This step is similar to the last step of the FSA.

Step 3 of the IFSA has the highest share in the computational time of the IFSA, which is improved from step  $i$  of the FSA. Let us consider two special cases of IFSA. If we divide the set of original scenarios,  $\Omega_{os}$ , into 1 cluster, i.e., the number of initial clusters is one, then  $\mathcal{C}_j = \Omega_{os}$ . If we divide  $\Omega_{os}$  into  $\|\Omega_{os}\|$  clusters, i.e., each cluster has only one element, then  $\mathcal{C} = \Omega_{os}$ . In other words, in these two special cases, the FSA is the same as the IFSA, i.e., the FSA is a special case of the IFSA. The setting of the number of initial clusters will be further discussed in Section V-B.

### E. Computational Complexity of Different Methods

The computational complexity of a method can be expressed by giving the total number of floating-point operations or flops required by the method [26]. The flops required by FSA is [12]:

$$f_{N_{os}}(N_{rs}) = 2N_{rs}^3/3 - N_{rs}^2(2N_{os} + 1) + N_{rs}(2N_{os}^2 + 2N_{os} + 1/3), \quad (14)$$

where  $N_{os}$  ( $N_{rs}$ ) represents the number of original scenarios (reduced scenarios). Equation (14) can be written as  $O(N_{os}^2 N_{rs})$  using the big  $O$  notation if  $N_{os}$  is much larger than  $N_{rs}$ .

The computational complexity of the K-means can be expressed as  $O(N_{os} N_{rs} n_d n_i)$ , where  $n_d$  is the dimension of a data in a scenario and  $n_i$  is the total number of iterations.

The computational complexity of the IFSA can be represented as  $O(N_{os}(k + N_{os}/k))$ , where  $k$  is the number of initial clusters. Then, it can be known that the computational load of the FSA is about  $N_{os} N_{rs} / (k + N_{os}/k)$  times as much as that of the IFSA. That is, when the number of original and reduced scenarios is large, the IFSA requires significantly less computational burden than the FSA.

## IV. STOCHASTIC TRANSMISSION EXPANSION PLANNING

### A. Stochastic Programming Model for TEP

The stochastic programming model for the TEP problem is given in (15a)-(15k), which is a nonlinear DC model as mentioned in [22]. The objective function consists of the generation cost, the wind spillage penalty cost, the load loss

penalty cost, and the investment cost for constructing the new lines as given in (15a). Constraint (15b) represents the power balance at each bus. Constraints (15c) and (15d) represent the capacity limits of each line. Constraints (15e), (15g), and (15i) represent the upper limits of the load loss, generation, and wind power curtailment, respectively. Constraints (15f), (15h), and (15j) represent that the load loss, generation, and wind power curtailment are nonnegative, respectively. Constraint (15k) represents the upper and lower limits of the new lines. Note that in the rest of the paper,  $\kappa$  is used to represent a scenario in the reduced set,  $\Omega_{rs}$ , obtained by a scenario reduction method.

Minimize:  $\sum_{\kappa \in \mathcal{C}} hp^\kappa (\rho_g^T g^\kappa + \rho_s^T S^\kappa + \rho_r^T r^\kappa) + \sum_{l \in \Omega} c_l n_l$  (15a)

$$\text{s.t. } \sum_{l=1}^L \left( M_{i,l} \tilde{\gamma}_l (n_l^0 + n_l) (\theta_{l,f}^\kappa - \theta_{l,t}^\kappa) \right) + g_i^\kappa + \bar{W}_i^\kappa$$

$$-S_i^\kappa + r_i^\kappa - d_i^\kappa = 0 : \lambda_{1,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15b)$$

$$\theta_{l,f}^\kappa - \theta_{l,t}^\kappa - f_l^{max} / \tilde{\gamma}_l \leq 0 : \mu_{1,l}^\kappa, \forall l \in \Omega_1, \forall \kappa \in \Omega_{rs} \quad (15c)$$

$$-(\theta_{l,f}^\kappa - \theta_{l,t}^\kappa) - f_l^{max} / \tilde{\gamma}_l \leq 0 : \mu_{2,l}^\kappa, \forall l \in \Omega_1, \forall \kappa \in \Omega_{rs} \quad (15d)$$

$$r_i^\kappa - d_i^\kappa \leq 0 : \mu_{3,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15e)$$

$$-r_i^\kappa \leq 0 : \mu_{4,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15f)$$

$$g_i^\kappa - \bar{g}_i \leq 0 : \mu_{5,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15g)$$

$$-g_i^\kappa \leq 0 : \mu_{6,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15h)$$

$$S_i^\kappa - \bar{W}_i^\kappa \leq 0 : \mu_{7,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15i)$$

$$-S_i^\kappa \leq 0 : \mu_{8,i}^\kappa, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (15j)$$

$$0 \leq n_l \leq \bar{n}_l, \forall l \in \Omega. \quad (15k)$$

where superscript  $T$  represents the transpose of a vector and subscript  $l$  represents the index of both existing and candidate lines. Note that  $\tilde{\gamma}_l (n_l^0 + n_l) (\theta_{l,f}^\kappa - \theta_{l,t}^\kappa)$  in (15b) represents the power flow of line  $l$  and that  $(\bar{W}_i^\kappa - S_i^\kappa)$  in (15b) represents the scheduled wind power generation at bus  $i$ .

### B. Benders Decomposition Algorithm

The STEP problem is solved by the Benders decomposition algorithm [27]. In the Benders decomposition algorithm for the STEP problem, the  $\kappa$ th slave problem can be defined as

$$\text{Minimize: } \varphi^\kappa = hp^\kappa (\rho_g^T g^\kappa + \rho_s^T S^\kappa + \rho_r^T r^\kappa) \quad (16a)$$

$$\text{s.t. (15b)-(15j)}$$

$$n_l - n_l^\xi = 0 : \lambda_{2,l}^\kappa, \forall l \in \Omega, \quad (16b)$$

and the master problem can be defined as

$$\text{Minimize: } \sum_{l \in \Omega} c_l n_l + \beta \quad (17a)$$

$$\text{s.t. } 0 \leq n_l \leq \bar{n}_l, \forall l \in \Omega \quad (17b)$$

$$\sum_{\kappa \in \mathcal{C}} \varphi^{\kappa,\xi} + \sum_{\kappa \in \mathcal{C}} \sum_{l \in \Omega} hp^\kappa \lambda_{2,l}^{\kappa,\xi} (n_l - n_l^\xi) \leq \beta, \forall \xi. \quad (17c)$$

In each iteration  $\xi$ , the solution  $n_l$  of the master problem is used as  $n_l^\xi$  in (16b) of the slave problem; and the Lagrange multiplier  $\lambda_{2,l}^\kappa$  associated to (16b) and the objective value  $\varphi^\kappa$  of the slave problem are respectively used as  $\lambda_{2,l}^{\kappa,\xi}$  and  $\varphi^{\kappa,\xi}$  in (17c).

In each iteration, after solving the master problem, the lower bound is calculated using

$$Z_{\text{down}}^\xi = \sum_{l \in \Omega} c_l n_l^\xi + \beta^\xi, \quad (18a)$$

and also after solving all the  $\kappa$  slave problems, the upper bound is calculated using

$$Z_{\text{up}}^\xi = \sum_{l \in \Omega} c_l n_l^\xi + \sum_{\kappa \in \mathcal{C}} \varphi^{\kappa,\xi}. \quad (18b)$$

The iterative process stops when the gap between the upper bound and the lower bound is below a predetermined value which is here set to  $0.01 \times \min(Z_{\text{up}}^\xi, Z_{\text{down}}^\xi)$  according to [5].

### C. Linearization Method for the Slave Problems

For the convenience of expression, minimizing (16a) constrained by (15b)-(15j) and (16b) are denoted as SNLP which represents a non-linear slave problem:

(SNLP) Minimize: (16a)

$$\text{s.t. (15b) - (15j), and (16b)}$$

Note that since  $n_l$  and  $(\theta_{l,f}^\kappa - \theta_{l,t}^\kappa)$  in constraint (15b) are variables, SNLP is a non-linear programming problem. By substituting (16b) into (15b), SNLP becomes a linear programming model:

(SLP) Minimize: (16a)

$$\text{s.t. } \sum_{l=1}^L \left( M_{i,l} \tilde{\gamma}_l (n_l^0 + n_l^\xi) (\theta_{l,f}^\kappa - \theta_{l,t}^\kappa) \right) + g_i^\kappa + \bar{W}_i^\kappa$$

$$-S_i^\kappa + r_i^\kappa - d_i^\kappa = 0 : \lambda_{1,i}^{\kappa,\xi}, \forall i \in \Omega_B, \forall \kappa \in \Omega_{rs} \quad (19)$$

$$(15c) - (15j).$$

For the convenience of expression, minimizing (16a) constrained by (19) and (15c)-(15j), are denoted as SLP which represents a linear slave problem. Note that the SLP is equivalent to the SNLP. Let  $\mathbf{x}^* = \{\mathbf{r}^{\kappa,*}, \mathbf{g}^{\kappa,*}, \mathbf{S}^{\kappa,*}, \boldsymbol{\theta}^{\kappa,*}\}$  be the optimal solution of the SLP. Then,  $\{\mathbf{x}^*, \mathbf{n}_L\}$ , where  $\mathbf{n}_L = \{n_1^\xi, n_2^\xi, \dots, n_L^\xi\}$ , is the optimal solution of the SNLP, which is proved in the second paragraph following (27c) in Appendix-A.

It has been proved that the SLP and the SNLP have the same solution. Therefore, we can solve the SLP to obtain the optimal solution of the SNLP instead of directly solving the SNLP. This is quite beneficial as the SLP is a linear problem and thereby is easier to be solved compared to the non-linear problem, SNLP. However, the SLP does not provide the Lagrangian multiplier associated with an important equality constraint in the SNLP as the SLP does not have that constraint. Note that this Lagrangian multiplier is required by the Benders decomposition algorithm. In this regard, a method to calculate the Lagrangian multiplier based on the solution obtained from solving the SLP is proposed.

To be more specific, the Lagrangian multiplier,  $\lambda_{2,l}^\kappa$ , associated with (16b) in SNLP, can not be directly obtained by solving SLP, while it is needed by (17c) in the master problem. In the following, a method to calculate  $\lambda_{2,l}^\kappa$  using the information obtained from solving the SLP is proposed.

The Lagrangian function of SNLP is:

$$\mathcal{L}^\kappa = hp^\kappa (\rho_g^T g^\kappa + \rho_s^T S^\kappa + \rho_r^T r^\kappa)$$

$$+ \sum_{l \in \Omega_1} \mu_{1,l}^\kappa (\theta_{l,f}^\kappa - \theta_{l,t}^\kappa - f_l^{max} / \tilde{\gamma}_l)$$

$$+ \sum_{l \in \Omega_1} \mu_{2,l}^\kappa (-\theta_{l,f}^\kappa + \theta_{l,t}^\kappa - f_l^{max} / \tilde{\gamma}_l)$$

$$+ \sum_{i=1}^N \mu_{3,i}^\kappa (r_i^\kappa - d_i^\kappa) + \sum_{i=1}^N \mu_{4,i}^\kappa (-r_i^\kappa)$$

$$+ \sum_{i=1}^N \mu_{5,i}^\kappa (g_i^\kappa - \bar{g}_i) + \sum_{i=1}^N \mu_{6,i}^\kappa (-g_i^\kappa)$$

$$+ \sum_{i=1}^N \mu_{7,i}^\kappa (S_i^\kappa - \bar{W}_i^\kappa) + \sum_{i=1}^N \mu_{8,i}^\kappa (-S_i^\kappa)$$

$$\begin{aligned}
& + \sum_{i=1}^N \lambda_{1,i}^{\kappa} [\sum_{l=1}^L (M_{i,l} \tilde{\gamma}_l (n_l^0 + n_l) (\theta_{l,f}^{\kappa} - \theta_{l,t}^{\kappa})) \\
& \quad + g_i^{\kappa} + \bar{W}_i^{\kappa} - S_i^{\kappa} + r_i^{\kappa} - d_i^{\kappa}] \\
& + \sum_{l=1}^L \lambda_{2,l}^{\kappa} (n_l - n_l^{\xi}), \quad \forall \kappa \in \Omega_{rs}. \quad (20)
\end{aligned}$$

According to the Karush–Kuhn–Tucker (KKT) conditions [27], the Lagrangian function with respect to  $n_l$  is equal to 0:

$$\frac{\partial L^{\kappa}}{\partial n_l} = \tilde{\gamma}_l (\theta_{l,f}^{\kappa} - \theta_{l,t}^{\kappa}) \sum_{i=1}^N (\lambda_{1,i}^{\kappa} M_{i,l}) + \lambda_{2,l}^{\kappa} = 0, \quad \forall l \in \Omega, \quad \forall \kappa \in \Omega_{rs}. \quad (21)$$

Equation (21) can be rewritten as:

$$\lambda_{2,l}^{\kappa} = -\tilde{\gamma}_l (\theta_{l,f}^{\kappa} - \theta_{l,t}^{\kappa}) \sum_{i=1}^N (\lambda_{1,i}^{\kappa} M_{i,l}), \quad \forall l \in \Omega, \quad \forall \kappa \in \Omega_{rs}. \quad (22)$$

As mentioned above, in the optimal solutions of both SNLP and SLP,  $\theta^{\kappa^*}$  is the same. Besides,  $\lambda_{1,i}^{\kappa^*}$  is equal to  $\lambda_{1,i}^{\kappa}$ , which is proved in the Appendix. Thus,  $\lambda_{2,l}^{\kappa^*}$ , which is associated with (16b) in SNLP, can be calculated using the optimal solution and the Lagrangian multipliers obtained from solving SLP:

$$\lambda_{2,l}^{\kappa^*} = -\tilde{\gamma}_l (\theta_{l,f}^{\kappa^*} - \theta_{l,t}^{\kappa^*}) \sum_{i=1}^N (\lambda_{1,i}^{\kappa^*} M_{i,l}), \quad \forall l \in \Omega, \quad \forall \kappa \in \Omega_{rs} \quad (23)$$

In this paper, the stochastic programming, where the uncertainty is represented by scenarios, is used to model the TEP problem and thereafter solved by a decomposition method. Although only the economy criterion is considered in the TEP, this framework can be extended to address other criteria such as stability [28], reliability [29], etc., by either adding these criteria as a constraint [30] or as another objective [31], which is an important topic [32] and will be our future work.

## V. SIMULATION

### A. Three Test Systems

To verify the effectiveness of the proposed method, the STEP problem is solved for three different test systems. The first one is an IEEE 24-bus system with 41 right-of-ways [8]. Using the same way as [8], the original generation capacity and load are multiplied by three to cause congestion in the system. Besides, two wind farms are connected to buses 1 and 15.

The second test system is an IEEE 300-bus system with 409 right-of-ways [33]. The original generation capacity and load are multiplied by 1.4 in order to cause congestion in the system. Three wind farms are connected to buses 1, 166, and 191. The third test system is a 2383-bus system with 2886 right-of-ways [33]. Three wind farms are connected to buses 4, 1116, and 1810. For each system, a maximum of three lines can be added to each right-of-way. The capacity of each wind farm is set to 1,500 MW.

In each of the modified test systems, the load and wind power capacity are determined, i.e., there is only one scenario with the probability of 1 and this is called the base case. For the convenience of expression, these determined load and wind power capacity are called as the base load and the base wind capacity, respectively. To obtain multiple scenarios, a load scenario factor,  $SF^{\kappa,L}$ , and a wind capacity scenario factor,  $SF_i^{\kappa,W}$ , at scenario  $\kappa$  are defined as follows:

$$SF^{\kappa,L} = L^{\kappa} / L^{max}, \quad \kappa \in \Omega_{rs}, \quad (24)$$

$$SF_i^{\kappa,W} = \bar{W}_i^{\kappa} / W^{max}, \quad \kappa \in \Omega_{rs}, \quad i = 1, 2, 3, \quad (25)$$

where  $L^{\kappa}$  ( $\bar{W}_i^{\kappa}$ ) is the value of load (wind capacity) in scenario  $\kappa$  which is obtained by a scenario reduction method, e.g., K-means or IFSA;  $L^{max}$  ( $W^{max}$ ) is the maximum value of  $L^{\kappa}$  ( $\bar{W}_i^{\kappa}$ ) for all  $\kappa \in \Omega_{rs}$ . The  $i$  in (25) is explained in the next paragraph.

The load data from 2000 to 2014 of Ontario, Canada [34] is used as the input data of scenario reduction to obtain  $L^{\kappa}$ . The wind speed data from three cities (Toronto, Wawa, and Peawanuck) in Ontario, in the same period from Environment Canada [35] is converted into wind power in the same way as [29]. The wind power is then used as the input data of scenario reduction to obtain  $\bar{W}_i^{\kappa}$ , where Toronto, Wawa, and Peawanuck are associated with  $i = 1, 2$ , and 3, respectively. For the 24-bus system,  $i = 1, 2$ , and for the other two systems,  $i = 1, 2$ , and 3.

Then, the load at each bus (the wind capacity of the  $i$ th wind farm) of each system in scenario  $\kappa$  is set as the base load (base wind capacity) multiplied by  $SF^{\kappa,L}$  ( $SF_i^{\kappa,W}$ ).

The STEP is solved using Matlab [36] on a PC with Intel Core i7-2770 3.40 GHz CPU and 16 GB RAM.

### B. Comparison Between Different Scenario Reduction Methods

FSA is implemented in the General Algebraic Modeling System (GAMS) [37] as it has a built-in function for FSA. The time consumed by FSA, K-means, and IFSA to reduce 13,104, 40,000, and 131,040 scenarios to 100 and 500 scenarios is tabulated in Table I. Note that when the number of original scenarios is 131,040, FSA is too time consuming to obtain a solution but K-means and IFSA can still obtain the solutions within 600 seconds. From Table I, it can be known that FSA (K-means) consumes the longest (shortest) time and that IFSA consumes a significantly shorter time than FSA and is capable of reducing a very large set of scenarios in a short time.

TABLE I  
TIME CONSUMED BY FSA, K-MEANS, AND IFSA FOR REDUCING DIFFERENT NUMBERS OF ORIGINAL SCENARIOS TO 100 AND 500 SCENARIOS.

Method	No. of reduced scenarios	No. of original scenarios		
		13,104	40,000	131,040
FSA	100	310 s	6,163 s	---
K-means		1.2 s	3.1 s	12.4 s
IFSA		3.7 s	21.4 s	164.7 s
FSA	500	1,411 s	24,493 s	---
K-means		13.8 s	45.6 s	109.3 s
IFSA		17.1 s	99.3 s	571.5 s

The number of initial clusters is an important parameter for the IFSA. The time consumption of the IFSA when using different numbers of initial clusters to reduce 13,104 and 40,000 scenarios to 100 scenarios is plotted in sub-figure a) of Figs. 1 and 2, respectively. Figs. 1 and 2 show that when the number of initial clusters is small, e.g., less than 10, the time consumption of IFSA is large and that the time consumed by the IFSA slowly increases when the number of initial clusters increases from 200 to 1000. This can be explained via the expression of the computational burden given in Section III-E, which indicates that the IFSA has the minimal computational burden when the number of initial clusters is equal to the square root of the number of original scenarios, which is equal to 114 and 200 in Figs. 1 and 2, respectively.

To compare the accuracy of the FSA, K-means, and IFSA, the space distance (the details can be found in [14]) between the original set and the reduced set obtained by the FSA, K-means,

and IFSA are shown in sub-figure b) of Figs. 1 and 2. Note that the lines associated with the K-means and FSA are straight. The reason is that the number of initial clusters is a parameter of the IFSA but not a parameter of the K-means and FSA.

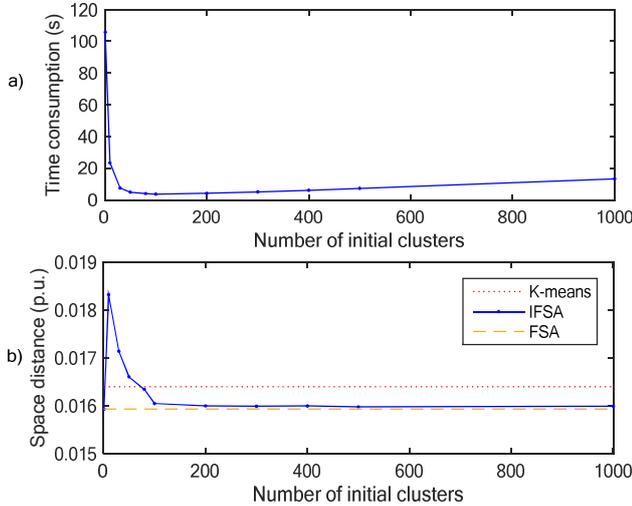


Fig. 1. a) Time consumed by IFSA in reducing 13,104 scenarios to 100 scenarios. b) Space distance between the original set of scenarios and the reduced set of scenarios obtained by K-means, IFSA, and FSA, respectively.

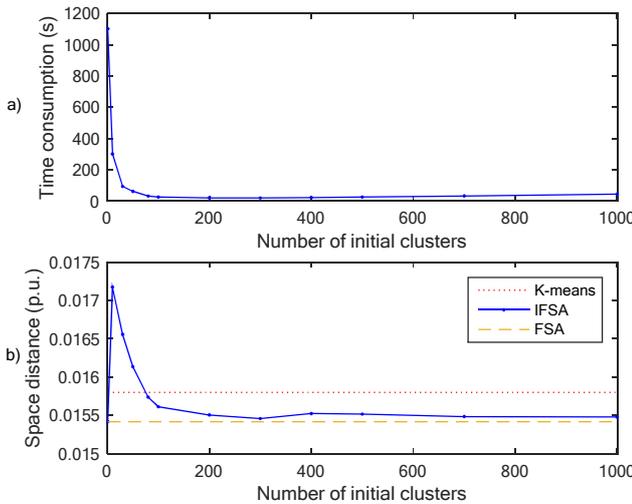


Fig. 2. a) Time consumed by IFSA in reducing 40,000 scenarios to 100 scenarios. b) Space distance between the original set of scenarios and the reduced set of scenarios obtained by K-means, IFSA, and FSA, respectively.

Sub-figure b) in both figures show that the space distance associated with the K-means (FSA) is the largest (smallest). In these two figures, the space distance associated with the K-means is 3.14% and 2.46% higher than that associated with the FSA, respectively. When the number of initial clusters is larger than a certain number, the space distance associated with the IFSA is quite close to that associated with the FSA (the relative error between them is 0.4% to 0.5%). The impacts of the scenario reduction accuracy associated with the K-means and the IFSA on the total cost of the STEP will be further investigated in Section V-E. As mentioned at the end of Section III-D, the IFSA has the same accuracy as the FSA when the

number of initial clusters is set to 1, which is verified in sub-figure b) in both Figs. 1 and 2. That is, the first point of the curve associated with the IFSA has the same space distance as that associated with the FSA.

It can be seen from sub-figure b) in both Figs. 1 and 2 that there is a knee point in the line associated with IFSA, so that the space distance decreases steeply (slightly) as the number of initial clusters increases on the left (right) side of the knee point. It is suggested that the number of initial clusters can be set at the knee point. For example, the number of initial clusters can be set to 200 and 300 in sub-figure b) of Figs. 1 and 2, respectively.

Now, we can come to the conclusion that the IFSA consumes a significantly shorter time than the FSA while the former is almost as accurate as the latter when the number of initial clusters is properly set.

### C. Deterministic Transmission Expansion Planning

For the purpose of demonstration and verification, the proposed linearization method is used to solve a deterministic TEP on the IEEE 24-bus system where the parameters are set to be the same as [23]. The total cost of the solution obtained is \$152,000,000 and the investment plan is  $n_{6-10} = 1$ ,  $n_{7-8} = 2$ ,  $n_{10-12} = 1$ , and  $n_{14-16} = 1$ , where the subscripts represent the bus numbers in the two ends of a right-of-way. These results are the same as those presented in [23].

### D. Time Consumption of Different Methods in Solving the STEP Using Different Test Systems

To verify the effectiveness of the linearization method proposed for the DC model used in the slave problems, three methods are implemented and compared with each other.

- Method 1: The slave problems in the Benders decomposition algorithm are directly solved as non-linear programming problems (i.e., **solving SNLP**).
- Method 2: The slave problems in the Benders decomposition algorithm are solved as mixed binary linear problems (i.e., **solving SBLP** given in Appendix-B). SBLP is a disjunctive model as mentioned in [22].
- Method 3: The slave problems in the Benders decomposition algorithm are solved as linear programming problems. That is, **solving SLP** and then using (23) to calculate the Lagrangian multiplier associated with (16b), which is the procedure of the proposed linearization method.

The number of variables and constraints in SBLP is more than those in SLP, especially in large-scale systems. The numbers of variables, inequality constraints, and equality constraints in terms of  $N$ ,  $|\Omega_0|$ ,  $|\Omega_1|$  and  $|\Omega_N|$  for SBLP and SLP are given in Table II, where  $|\Omega_0| \leq L$ ,  $|\Omega_N| \leq \bar{n}_l L$ , and  $|\Omega_1| \leq L$ . Note that (15e)-(15j) in SLP are handled as upper and lower bounds and that (34b) is substituted into (34a) and (34c) in SBLP. The numbers of variables and inequality constraints for SBLP and SLP for the three test systems are plotted in Fig. 3, where ' $n_1 E + n_2$ ' represents ' $n_1 \times 10^{n_2}$ '. Table II and Fig. 3 show that SBLP has many more variables and inequality constraints than SLP, especially when  $|\Omega_N|$  is large.

The SNLP model is nonlinear while the SBLP and SLP models are linear. The time consumption for solving each of these three models in each of the three test systems is tabulated

in Table III, where ‘s’ represents seconds. Table III shows that solving SNLP consumes much more time than solving SBLP or SLP and that solving SBLP consumes more time than solving SLP (about 2.2 to 5.5 times).

TABLE II  
NUMBER OF VARIABLES, INEQUALITY CONSTRAINTS, AND EQUALITY CONSTRAINTS IN SBLP AND SLP.

	SBLP	SLP
No. of variables	$4N +  \Omega_N $	$4N$
No. of inequality constraints	$2 \Omega_0  + 4 \Omega_N $	$2 \Omega_1 $
No. of equality constraints	$N$	$N$

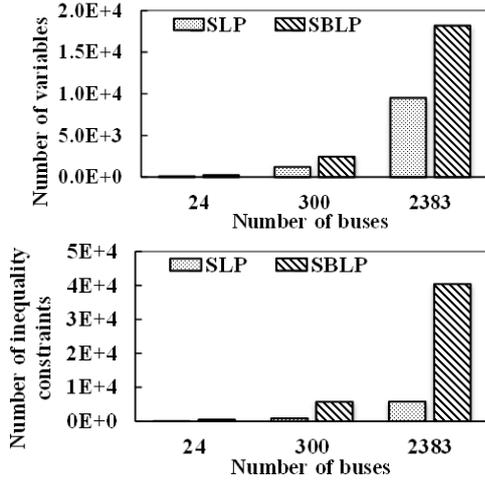


Fig. 3. The numbers of variables and inequality constraints in SLP and SBLP for each of the three test systems, respectively.

TABLE III  
TIME CONSUMPTION FOR SOLVING MODELS SNLP, SBLP AND SLP IN EACH OF THE THREE TEST SYSTEMS.

Test problem	SNLP	SBLP	SLP
24-bus	3 s	0.0026 s	0.0012 s
300-bus	33.4 s	0.0904 s	0.0164 s
2383-bus	--	5.7 s	1.4 s

Methods 1-3 are different in solving the slave problems but are the same in solving the master problem. Thus, only the time consumed by the three methods to solve the slave problems is compared here. The time consumed by the three methods to solve all the slave problems in each of the three test systems is given in Table IV, where ‘s’ represents seconds and ‘d’ represents days. It can be seen from the table that Method 1 is too time consuming to be used for solving the STEP problem. Besides, Method 2 consumes 343 seconds, 4810 seconds, and 3.5 days more than Method 3 for the three test systems, respectively. That is, the advantage of Method 3 over Method 2 becomes more obvious as the system gets larger.

TABLE IV  
TOTAL TIME CONSUMED BY THE THREE METHODS TO SOLVE ALL THE SLAVE PROBLEMS IN EACH OF THE THREE TEST SYSTEMS, RESPECTIVELY.

Test problem	Method 1	Method 2	Method 3
24-bus	735,000 s (8.5 d)	637 s	294 s
300-bus	2,171,000 s (25.1 d)	5,876 s	1,066 s
2383-bus	--	399,000 s (4.6 d)	98,000 s (1.1 d)

### E. Number of Reduced Scenarios

The number of reduced scenarios is an important parameter for the STEP problem. Obviously, as the number of reduced scenarios increases, the reduced scenarios can better represent the original scenarios and consequently a more accurate solution can be obtained for the STEP. Various numbers of reduced scenarios, obtained by both the K-means clustering method and the IFSA, are used for solving the STEP problem and the associated total costs for the 24-bus system as a function of the number of reduced scenarios are plotted in Fig. 4.

Fig. 4 shows that the total cost varies significantly with the increase of the number of reduced scenarios in the first 5 points while the variation of the total cost is small in the last 4 points. The reason is that the reduced scenarios cannot (can) accurately represent the original scenarios when the number of reduced scenarios is smaller (larger) than a certain number. Thereby, the total costs of the STEP should converge to the optimal result when the number of reduced scenarios is larger than this certain number (equal to 10,000 in Fig. 4), which is supported in Fig. 4. This certain number is suggested to be the proper setting of the number of reduced scenarios.

That is, the number of reduced scenarios needs to be relatively large in order to obtain an accurate result. However, this requires a long computational time for both the scenario reduction methods and the solution methods to solve the STEP, which makes the IFSA and the linearization method proposed in Sections III and IV, respectively, more valuable.

It can be seen from Fig. 4 that the total cost associated with the K-means is larger than that associated with the IFSA. Actually, when the number of reduced scenarios is equal to 15,000, the total costs associated with the K-means and the IFSA are  $1.49 \times 10^9$  \$ and  $1.37 \times 10^9$  \$, respectively, and the former is 8.76% (120 M\$) larger than the latter. This indicates the importance of the accuracy of scenario reduction methods for the STEP problem.

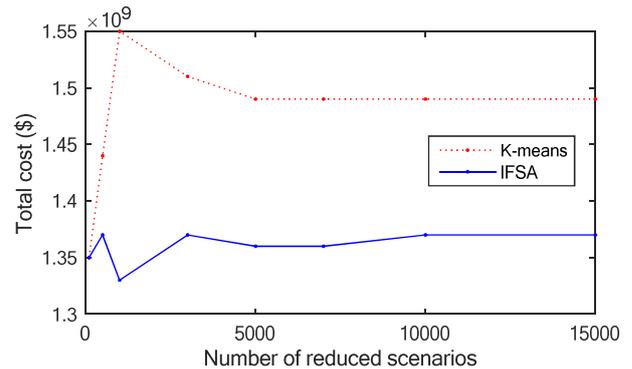


Fig. 4. Total cost of the STEP using various numbers of reduced scenarios.

### F. Cost of Load Loss

Another important parameter is the cost of load loss  $\rho_r$ . Various values of  $\rho_r$  are used in the STEP and the results are tabulated in Table V, where the weighted sum of load loss is calculated by  $\sum_{k \in C} p^k (\sum_{i \in \Omega_B} r_i^k)$ . It can be seen from Table V that when the cost of load loss is low (i.e.,  $\rho_r$  is from 10 to 100), only a few new lines are added. This is because the cost of load loss is lower than adding new lines and thereby there is much

load shedding. Table V also shows that as the cost of load loss increases, the weighted sum of load loss decreases.

The total cost and the weighted sum of load loss as a function of the cost of load loss are plotted in Fig. 5, where ‘1E+n’ represents ‘ $1 \times 10^n$ ’. Logarithmic scale is used to represent the total cost and the cost of load loss. It can be seen from the figure that the solid line, representing the total cost, increases steadily while the dashed line, representing the weighted sum of load loss, has an obvious knee point when the cost of load loss is equal to  $1 \times 10^3$ . This indicates that the cost of load loss can be set to  $1 \times 10^3$  to maintain a tradeoff between the total cost and the weighted sum of load loss.

TABLE V  
THE RESULTS OF THE STEP IN THE 24-BUS SYSTEM WHERE VARIOUS VALUES OF THE COST OF LOAD LOSS ARE USED.

$\rho_r$	10	100	$1 \times 10^3$
Total cost (\$)	$6.95 \times 10^8$	$9.75 \times 10^8$	$1.36 \times 10^9$
No. of new lines	3	9	15
Weighted sum of load loss (MW)	42.1749	1.1546	0.2532
$\rho_r$	$1 \times 10^4$	$1 \times 10^5$	$1 \times 10^6$
Total cost (\$)	$2.77 \times 10^9$	$1.49 \times 10^{10}$	$1.34 \times 10^{11}$
No. of new lines	19	27	37
Weighted sum of load loss (MW)	0.1613	0.1512	0.1505

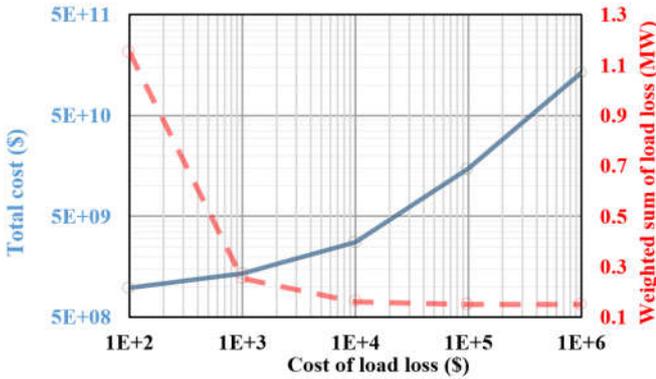


Fig. 5. Total cost and weighted sum of load loss as a function of the cost of load loss.

## VI. CONCLUSION

In this paper, a new scenario reduction method, the IFSA, has been proposed. The reduced scenarios are then used in the stochastic programming model for the TEP. The STEP is decomposed into a master problem and multiple slave problems by using the Benders decomposition algorithm. A linearization method has been proposed to convert the slave problems to linear programming models and calculate the Lagrangian multipliers needed in the master problem.

For the purpose of comparison, the K-means clustering method, the FSA, and the IFSA are used for scenario reduction of two medium and one large sets of scenarios. The simulation results have shown that the K-means consumes the shortest time but its accuracy is about 2.5% to 3% lower than the FSA and the IFSA, that the FSA is the most accurate but quite time consuming, and that the IFSA is almost as accurate as the FSA but consumes a significantly shorter time than the FSA. The simulation results have also shown that the FSA is too time consuming to reduce a large number of scenarios while the K-means and the IFSA can reduce a large number of scenarios in

a short time. Besides, the total cost of the STEP using the reduced scenarios obtained by the K-means is much higher than that associated with the IFSA, which indicates the necessity of using the IFSA instead of the K-means.

The STEP problem has been solved in the modified 24-, 300-, and 2383-bus test systems. The simulation results have shown that the linearization method proposed for the DC models of the slave problems can remarkably reduce their complexity and solution time and that it can cause the slave problems to be solved much faster (especially for large-scale problems) than an existing method that uses disjunctive models of the slave problems.

The number of reduced scenarios has been investigated and the simulation results show that a relatively large number of the reduced scenarios are necessary in order to obtain an accurate solution for the STEP. The optimal value for the cost of load loss has also been investigated in order to maintain a tradeoff between the total cost and the load loss.

## APPENDIX

### A. Proof of SNLP and SLP Having the Same Lagrangian Multipliers

SNLP can be represented as:

$$(NLP) \quad \text{Minimize: } f(\mathbf{x}_1) \quad (26a)$$

$$\text{s.t. } \mathbf{g}(\mathbf{x}_1) \leq 0 : \boldsymbol{\mu}_1 \quad (26b)$$

$$\mathbf{h}(\mathbf{x}_1^T \mathbf{x}_2) = 0 : \boldsymbol{\lambda}_1 \quad (26c)$$

$$\mathbf{x}_2 - \mathbf{a}_0 = 0 : \boldsymbol{\lambda}_2, \quad (26d)$$

where  $f$  and  $\mathbf{g}$  are linear functions of  $\mathbf{x}_1$ ;  $\mathbf{h}$  is a linear function of  $\mathbf{x}_1^T \mathbf{x}_2$ ;  $\mathbf{a}_0$  is a vector of parameters; and  $\boldsymbol{\mu}_1$ ,  $\boldsymbol{\lambda}_1$ , and  $\boldsymbol{\lambda}_2$  are Lagrangian multipliers associated with (26b)-(26d), respectively. Note that (26b) represents (15c)-(15j); (26c) represents (15b); and (26d) represents (16b). For the convenience of expression, (26a)-(26d) are denoted as NLP, which is a non-linear programming problem.

SLP can be represented as:

$$(LP) \quad \text{Minimize: } f(\mathbf{x}_1) \quad (27a)$$

$$\text{s.t. } \mathbf{g}(\mathbf{x}_1) \leq 0 : \boldsymbol{\mu}'_1 \quad (27b)$$

$$\mathbf{h}(\mathbf{x}_1^T \mathbf{a}_0) = 0 : \boldsymbol{\lambda}'_1, \quad (27c)$$

where  $\mathbf{h}$  is a linear function of  $\mathbf{x}_1^T \mathbf{a}_0$ ; and  $\boldsymbol{\mu}'_1$  and  $\boldsymbol{\lambda}'_1$  are Lagrangian multipliers associated with (27b)-(27c), respectively. Note that (27b) represents (15c)-(15j) and (27c) represents (19). For the convenience of expression, (27a)-(27c) are denoted as LP, which is a linear programming problem.

Denote the optimal solution of LP as  $\mathbf{x}_1^*$ . Then, the optimal solution of NLP is  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  where  $\mathbf{x}_2^* = \mathbf{a}_0$  which is proved as follows. Suppose  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  is not the optimal solution of NLP. Then, there exist another feasible solution  $(\mathbf{x}_1^0, \mathbf{x}_2^*)$  such that  $f(\mathbf{x}_1^0) < f(\mathbf{x}_1^*)$ , which contradicts with the assumption that  $\mathbf{x}_1^*$  is the optimal solution of LP. Thus, it has been proved that  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  is the optimal solution of NLP. Similarly, it can be proved that if the optimal solution of NLP is denoted as  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$ , then  $\mathbf{x}_1^*$  is the optimal solution of LP.

Now we want to prove that  $\boldsymbol{\mu}_1 = \boldsymbol{\mu}'_1$  and  $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}'_1$ .

According to [27], the dual problem of LP is:

$$\text{Maximize: } \phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1) \quad (28)$$

$$\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1; \boldsymbol{\mu}'_1 \geq 0$$

where

$$\phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1) = \text{Infimum}_{\mathbf{x}_1} \{f(\mathbf{x}_1) + \boldsymbol{\mu}'_1 \mathbf{g}(\mathbf{x}_1) + \boldsymbol{\lambda}'_1 \mathbf{h}(\mathbf{x}_1^T \mathbf{a}_0)\} \quad (29)$$

Denote the optimal value of (28) as  $d^* = \phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1)$  where  $(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1)$  are the optimal solution of (28). Denote the optimal value of LP as  $f^* = f(\mathbf{x}_1^*)$  where  $\mathbf{x}_1^*$  is the optimal solution of LP. According to [26], strong duality holds for any linear programming provided the primal problem is feasible, i.e.,  $d^* = f^*$ , which can be written as:

$$d^* = \phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1) = \max_{\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1; \boldsymbol{\mu}'_1 \geq 0} \phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1)$$

$$= \max_{\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1; \boldsymbol{\mu}'_1 \geq 0} [f(\mathbf{x}_1^*) + \boldsymbol{\mu}'_1 \mathbf{g}(\mathbf{x}_1^*) + \boldsymbol{\lambda}'_1 \mathbf{h}(\mathbf{x}_1^{*T} \mathbf{a}_0)]. \quad (30)$$

According to [27], the dual problem of NLP is:

$$\text{Maximize : } \phi_1(\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) \quad (31)$$

$$\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \boldsymbol{\mu}_1 \geq 0$$

where

$$\phi_1(\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \text{Infimum}_{\mathbf{x}_1, \mathbf{x}_2} \{f(\mathbf{x}_1) + \boldsymbol{\mu}_1 \mathbf{g}(\mathbf{x}_1) + \boldsymbol{\lambda}_1 \mathbf{h}(\mathbf{x}_1^T \mathbf{x}_2) + \boldsymbol{\lambda}_2 (\mathbf{x}_2 - \mathbf{a}_0)\} \quad (32)$$

As mentioned above,  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  where  $\mathbf{x}_2^* = \mathbf{a}_0$  is the optimal solution of NLP and its optimal value is  $f^*$ . Substituting  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  into (32), (31) becomes:

$$\text{Maximize : } \phi_1(\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \max_{\boldsymbol{\mu}_1, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2; \boldsymbol{\mu}_1 \geq 0} [f(\mathbf{x}_1^*) + \boldsymbol{\mu}_1 \mathbf{g}(\mathbf{x}_1^*) + \boldsymbol{\lambda}_1 \mathbf{h}(\mathbf{x}_1^{*T} \mathbf{a}_0) + \boldsymbol{\lambda}_2 (\mathbf{a}_0 - \mathbf{a}_0)] \quad (33)$$

Denote the optimal solution of (33) as  $(\boldsymbol{\mu}_1^*, \boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^*)$ . Note that the right hand side of (33) is equivalent to the second line of (30) as the last term in the right hand side of (33) is 0. Thus, the first two elements of the optimal solution of (33) are equal to the optimal solution of (30), i.e.,  $\boldsymbol{\mu}_1^* = \boldsymbol{\mu}'_1$  and  $\boldsymbol{\lambda}_1^* = \boldsymbol{\lambda}'_1$ . Besides, the optimal value of (33) is equal to the optimal value of (30):  $\phi_1(\boldsymbol{\mu}_1^*, \boldsymbol{\lambda}_1^*, \boldsymbol{\lambda}_2^*) = \phi_2(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1) = d^* = f^*$ . That is, when substituting  $(\mathbf{x}_1^*, \mathbf{x}_2^*)$  into (32), the value of the dual problem (31) is  $d^*$  which is equal to the optimal value of primal problem NLP. This indicates that the strong duality holds and that the optimal solution of (33) is the same as the optimal solution of (31). Note that the optimal solution of (30) is the same as the optimal solution of (28). Thus, the first two elements,  $(\boldsymbol{\mu}_1^*, \boldsymbol{\lambda}_1^*)$ , of the optimal solution of (31) are equal to the optimal solution of (28),  $(\boldsymbol{\mu}'_1, \boldsymbol{\lambda}'_1)$ . Note that the Lagrangian multipliers of the primal problems SNLP and SLP are equal to the values of the optimal solutions of the dual problems (31) and (28), respectively. Thus, it has been proved that the Lagrangian multipliers of (26b)-(26c) are equal to those of (27b)-(27c), respectively, i.e.,  $\boldsymbol{\mu}_1 = \boldsymbol{\mu}'_1$  and  $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}'_1$ .

*B. Mixed Binary Linear Programming Model for Slave Problems*

(SBLP) Minimize: (16a)

$$\text{s.t. } \sum_{i \in \Omega_0} M_{i,l} f_l^K + \sum_{n \in \Omega_N} M_{i,n} f_n^K + g_i^K + \bar{W}_i^K - S_i^K$$

$$+ r_i^K - d_i^K = 0 : \lambda_{1,i}^K, \forall i \in \Omega_B, \forall K \in \mathcal{C} \quad (34a)$$

$$f_l^K - \tilde{\gamma}_l n_l^0 (\theta_{l,f}^K - \theta_{l,t}^K) = 0, \forall l \in \Omega_0, \forall K \in \mathcal{C} \quad (34b)$$

$$|f_l^K| \leq n_l^0 f_l^{max}, \forall l \in \Omega_0, \forall K \in \mathcal{C} \quad (34c)$$

$$-B_M(1 - x_n) \leq f_n^K - \tilde{\gamma}_n (\theta_{n,f}^K - \theta_{n,t}^K) \leq B_M(1 - x_n), \forall n \in \Omega_N, \forall K \in \mathcal{C} \quad (34d)$$

$$-x_n f_n^{max} \leq f_n^K \leq x_n f_n^{max}, \forall n \in \Omega_N, \forall K \in \mathcal{C} \quad (34e)$$

$$x_n \in \{0,1\}, \forall n \in \Omega_N \quad (34f)$$

This is a mixed binary linear programming model for slave problems. For the convenience of expression, this optimization model is denoted as SBLP which represents a linear slave problem with binary variables.

## REFERENCES

- [1] G. Latorre, R.D. Cruz, J.M. Areiza, and A. Villegas, "Classification of publications and models on transmission expansion planning", *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 938-946, 2003.
- [2] C.W. Lee, S.K.K. Ng, J. Zhong, and F.F. Wu, "Transmission expansion planning from past to future," in *IEEE PES Power Syst. Conf. Exposition*, Atlanta, GA, USA, pp. 257-265, Oct. 2006.
- [3] Global wind energy council, *Global wind report – Annual market update 2015*, Available: <http://www.gwec.net>.
- [4] A.H. van der Weijde, and B.F. Hobbs, "The economics of planning electricity transmission to accommodate renewables: using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty", *Energy Economics*, vol. 34, no. 6, pp. 2089-2101, 2012.
- [5] H. Park, R. Baldick, and D.P. Morton, "A stochastic transmission planning model with dependent load and wind forecasts", *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3003-3011, 2015.
- [6] L. Baringo, and A.J. Conejo, "Transmission and wind power investment", *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 885-893, 2012.
- [7] B. Chen, J. Wang, L. Wang, Y. He, and Z. Wang, "Robust optimization for transmission expansion planning: minimax cost vs. minimax regret", *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3069-3077, 2014.
- [8] H. Yu, C.Y. Chung, K.P. Wong, and J.H. Zhang, "A chance constrained transmission network expansion planning method with consideration of load and wind farm uncertainties", *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1568-1576, 2009.
- [9] J. Choi, T. Tran, A.A. El-Keib, R. Thomas, H. Oh, and R. Billinton, "A method for transmission system expansion planning considering probabilistic reliability criteria", *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1606-1615, 2005.
- [10] H. Pandzic, Y. Dvorkin, T. Qiu, Y. Wang, and D.S. Kirschen, "Toward cost-efficient and reliable unit commitment under uncertainty", *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 970-982, 2016.
- [11] H. Heitsch, and W. Romisch, "A note on scenario reduction for two-stage stochastic programs", *Operations Res. Lett.*, vol. 35, no. 6, pp. 731-738, 2007.
- [12] H. Heitsch, and W. Romisch, "Scenario reduction algorithms in stochastic programming", *Comput. Optimization and Appl.*, vol. 24, no. 2-3, pp. 187-206, 2003.
- [13] J.M. Morales, S. Pineda, A.J. Conejo, and M. Carrion, "Scenario reduction for futures market trading in electricity markets", *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 878-888, 2009.
- [14] J. Li, F. Lan, and H. Wei, "A scenario optimal reduction method for wind power time series", *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1657-1658, 2016.
- [15] F.D. Munoz, B.F. Hobbs, and J.P. Watson, "New bounding and decomposition approaches for MILP investment problems: Multi-area transmission and generation planning under policy constraints", *European J. Operational Res.*, vol. 248, no. 3, pp. 888-898, 2016.
- [16] H. Heitsch, and W. Romisch, "Scenario tree modeling for multistage stochastic programs", *Mathematical Programming*, vol. 118, no. 2, pp. 371-406, 2009.
- [17] V.S. Pappala, I. Erlich, K. Rohrig, and J. Dobschinski, "A stochastic model for the optimal operation of a wind-thermal power system", *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 940-950, 2009.
- [18] H. Li, E. Haq, K. Abdul-Rahman, J. Wu, and P. Causgrove, "On-line voltage security assessment and control", *Int. Trans. Elect. Energy Syst.*, no. 24, pp.

- 1618-1631, 2014.
- [19] W. Wu, K. Y. Wang, B. Han, G. J. Li, X. C. Jiang, and M. L. Crow, "A versatile probability model of photovoltaic generation using pair copula construction", *IEEE Trans. Sustainable Energy*, Vol. 6, No. 4, pp. 1337-1345, 2015.
- [20] G. Díaz, P. G. Casielles, and J. Coto, "Simulation of spatially correlated wind power in small geographic areas—Sampling methods and evaluation", *Int. J. Elec. Power & Energy Syst.*, Vol. 63, pp. 513-522, 2014.
- [21] D. F. Cai, D. Y. Shi, and J. F. Chen, "Probabilistic load flow computation using Copula and Latin hypercube sampling", *IET Gener. Transm. Distrib.* Vol. 8, No. 9, pp. 1539-1549, 2014.
- [22] R. Romero, A. Monticelli, A. Garcia, and S. Haffner, "Test systems and mathematical models for transmission network expansion planning", *IEE Proceedings - Generation, Transmission and Distribution*, vol. 149, no. 1, pp. 27-36, 2002.
- [23] I. De J. Silva, M.J. Rider, R. Romero, and C.A.F. Murari, "Transmission network expansion planning considering uncertainty in demand", *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1565-1573, 2006.
- [24] H. Park, and R. Baldick, "Transmission planning under uncertainties of wind and load: sequential approximation approach", *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2395-2402, 2013.
- [25] D. Arthur, and S. Vassilvitskii, "K-means++: The advantages of careful seeding." *SODA '07*, Society for Industrial and Applied Mathematics, pp. 1027-1035, 2007.
- [26] S. Boyd, and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [27] A.J. Conejo, E. Castillo, R. Minguez, and R. Garcia-Bertrand, *Decomposition Techniques in Mathematical Programming*, Springer, 2006.
- [28] S. Robak, and K. Gryszpanowicz, "Rotor angle small signal stability assessment in transmission network expansion planning", *Elect. Power Syst. Res.*, vol. 128, pp. 144-150, 2015.
- [29] R. Billinton, and W. Wangdee, "Reliability-based transmission reinforcement planning associated with large-scale wind farms", *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 34-41, 2007.
- [30] P. Jirutitijaroen, and C. Singh, "Reliability constrained multi-area adequacy planning using stochastic programming with sample-average approximations", *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 504-513, 2008.
- [31] P. Maghouli, S.H. Hosseini, M.O. Buygi, and M. Shahidehpour, "A scenario-based multi-objective model for multi-stage transmission expansion planning", *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 470-478, 2011.
- [32] J. Quintero, H. Zhang, Y. Chakhchoukh, V. Vittal, and G.T. Heydt, "Next generation transmission expansion planning framework: models, tools, and educational opportunities", *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1911-1918, 2014.
- [33] R.D. Zimmerman, C.E. Murillo-Sanchez, and R.J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education", *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12-19, 2011.
- [34] Ontario's IESO, *IESO Historical Reports*. [Online]. Available: <http://www.ieso.ca/Pages/Power-Data/Data-Directory.aspx>.
- [35] Government of Canada, *Historical Climate Data*. [Online]. Available: [http://climate.weather.gc.ca/index\\_e.html#access=Opt](http://climate.weather.gc.ca/index_e.html#access=Opt).
- [36] MATLAB and Statistics Toolbox Release 2013a. The MathWorks, Inc., Natick, Massachusetts, United States.
- [37] GAMS Development Corporation, *General Algebraic Modeling System (GAMS) Release 24.2.1*. Washington, DC, USA, 2013.

**Junpeng Zhan** (M'16) received B.Eng. and Ph.D. degrees in electrical engineering from Zhejiang University, Hangzhou, China in 2009 and 2014, respectively.

He is currently a Postdoctoral Fellow in the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada. His current research interests include the integration of the energy storage systems, dynamic thermal rating and renewable electric energy sources in power systems.

**C. Y. Chung** (M'01-SM'07-F'16) received B.Eng. (with First Class Honors) and Ph.D. degrees in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, China, in 1995 and 1999, respectively.

He has worked for Powertech Labs, Inc., Surrey, BC, Canada; the University of Alberta, Edmonton, AB, Canada; and The Hong Kong Polytechnic University, China. He is currently a Professor, the NSERC/SaskPower (Senior) Industrial Research Chair in Smart Grid Technologies, and the SaskPower

Chair in Power Systems Engineering in the Department of Electrical and Computer Engineering at the University of Saskatchewan, Saskatoon, SK, Canada. His research interests include smart grid technologies, renewable energy, power system stability/control, planning and operation, computational intelligence applications, and power markets.

Dr. Chung is an Editor of *IEEE Transactions on Power Systems* and *IEEE Transactions on Sustainable Energy* and an Associate Editor of *IET Generation, Transmission, and Distribution*. He is also an IEEE PES Distinguished Lecturer and a Member-at-Large (Global Outreach) of the IEEE PES Governing Board.

**Alireza Zare** (S'14) received the B.Sc. degree in electrical engineering from Shahid Rajaei Teacher Training University, Tehran, Iran, in 2012, and the M.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2014. In September 2015, he joined the Department of Electrical and Computer Engineering at the University of Saskatchewan, Canada, where he is currently working toward the Ph.D. degree in electrical engineering.

His research interests include distribution system planning and operation, energy management of electric vehicles, renewable and distributed energy resources, and smart grid technologies.