Data-driven Power Flow Method Based on Exact Linear Regression Equations

Yanbo Chen, Chao Wu, and Junjian Qi

***Abstract*—Power flow (PF) is one of the most important calculations in power systems. The widely-used PF methods are the Newton-Raphson PF (NRPF) method and the fast-decoupled PF (FDPF) method. In smart grids, power generations and loads become intermittent and much more uncertain, and the topology also changes more frequently, which may result in significant state shifts and further make NRPF or FDPF difficult to converge. To address this problem, we propose a data-driven PF (DDPF) method based on historical/simulated data that includes an offline learning stage and an online computing stage. In the offline learning stage, a learning model is constructed based on the proposed exact linear regression equations, and then the proposed learning model is solved by the ridge regression (RR) method to suppress the effect of data collinearity. In online computing stage, the nonlinear iterative calculation is not needed. Simulation results demonstrate that the proposed DDPF method has no convergence problem and has much higher calculation efficiency than NRPF or FDPF while ensuring similar calculation accuracy.**

***Index Terms*—Data driven, exact linear regression equation, Fast-decoupled power flow, Newton-Raphson method.**

I. Introduction

POWER flow (PF) calculation, one of the most important calculations in power system, is widely-used in power system planning, operation, and control. The earliest proposed PF methods include Gauss-Seidel method and Newton-Raphson PF (NRPF) method [1]. NRPF has better convergence and higher computational efficiency than Gauss-Seidel method. Sparse techniques are also proposed to substantially improve the computational efficiency of PF calculation [2]. To further improve the computational efficiency of NRPF, the fast-decoupled PF (FDPF) method is proposed [3]. The above PF methods need to first construct the nonlinear PF equations based on the bus admittance matrix, and then apply iterative methods to solve them. These methods are called model-driven PF (MDPF) methods.

In the emerging smart grids, MDPF methods may have significant shortcomings. First, power generations and loads become intermittent and much more uncertain, and the topology also changes frequently, resulting in significant state shifts. This will make the widely-used NRPF or FDPF difficult to converge [4], as they usually use the PF result of the previous snapshot as initial guess. Second, the grid parameters have a certain degree of uncertainty. The grid parameters stored in the database of power dispatching control centers may differ from their actual values, which will negatively affect the performance of the MDPF methods.

In the smart grids, a large amount of historical/simulation data are available. Based on these data, data-driven methods are effective means to improve the calculation accuracy and efficiency of traditional model driven power system analysis methods [4]. Several data-driven power system analysis methods have been proposed, but few of them are data-driven PF (DDPF) methods. In [5], a novel model-free control (MFC) based emergency control scheme is presented. In [6], the support vector regression (SVR) model is used to train the mapping relationship between injection power and state variables by using historical data, which cannot handle PV buses. In [7], a DDPF method including a forward regression model and an inverse regression model is proposed. But the method in [7] cannot handle different topologies.

DDPF methods generally need to use historical or simulation data as sample data and construct a learning model to obtain the mapping relationship between the boundary conditions of PF calculation and state variables. Through in-depth analysis, we find that the difficulty of DDPF modeling is mostly attributed to the nonlinearity of the PF equations. The nonlinear PF equations may lead to the over-learning problems of the existing DDPF learning models and affect their calculation accuracy and computing efficiency. If the original nonlinear PF equations can be accurately transformed into linear equations, the calculation accuracy and the computing efficiency of the existing DDPF learning models may be improved. This motivates us to propose a novel DDPF method based on exact linear regression equations (ELREs).

II. Exact Linear PF Equations

The original nonlinear PF equations are as follows:

Manuscript received: October 12, 2020; revised: December 13, 2020; accepted: March 2, 2021. Date of CrossCheck: March 2, 2021. Date of online publication: June 4, 2021.

This work was supported in part by National Natural Science Foundation of China (No. 52077076) and in part by the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (No. LAPS202118).

Y. Chen (corresponding anthor) and C. Wu are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Electrical & Electronic Engineering, North China Electric Power University, Beijing 102206, China (e-mail: chenyanbo@ncepu. edu. cn; 1125027552@qq. com).

J. Qi is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA (e-mail: jqi8@stevens.edu).

DOI: 10.35833/MPCE.2020.000738

$$P_i = V_i \sum_{j \in k(i)} V_j (G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij}) \tag{1}$$

$$Q_i = V_i \sum_{j \in k(i)} V_j (G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij}) \tag{2}$$

where P_i and Q_i are the active and reactive power injections at bus i, respectively; $k(i)$ denotes all buses directly connected to bus i (including i); and $G_{ij}+\mathrm{j}B_{ij}$ is the element in the bus admittance matrix of node j and $G_{ii}+\mathrm{j}B_{ii}$ is the element in the bus admittance matrix of node i. Without the loss of generality, we assume bus 1 is the slack bus and the state variables (voltage magnitudes and phase angles) are $V_i \angle \theta_i (i = 2, 3, \ldots, N)$.

Base on nonlinear PF equations (1) and (2), MDPF methods can be implemented. As in [8], the nonlinear PF equations (1) and (2) can be transformed into the exact linear PF equations:

$$P_i = G_{ii} U_i + \sum_{j \in k(i), j \neq i} (G_{ij} R_{ij} + B_{ij} I_{ij}) \tag{3}$$

$$Q_i = -B_{ii} U_i + \sum_{j \in k(i), j \neq i} (G_{ij} I_{ij} - B_{ij} R_{ij}) \tag{4}$$

where R_{ij}, I_{ij}, and U_i are the auxiliary state variables, $R_{ij} = V_i V_j \cos\theta_{ij}$, $I_{ij} = V_i V_j \sin\theta_{ij}$, and $U_i = V_i^2$.

The following equation should be built for each PV bus:

$$U_{i,PV} = U_i \tag{5}$$

where $U_{i,PV} = V_{i,PV}^2$, and $V_{i,PV}$ is the given voltage magnitude of the i^{th} PV bus.

After the introduction of auxiliary state variables, the original nonlinear PF equations are accurately transformed into linear counterparts. For each PQ bus, linear PF equations (3) and (4) should be established. And for each PV bus, linear PF equations (3) and (5) should be established. The total number of linear PF equations is $2(N-1)$, while the total number of auxiliary state variables is $N-1+2b$, where b is the number of branches.

III. Proposed DDPF Method

A. Formulation of Exact Linear Regression Equations

Equations (3)-(5) can be further uniformly expressed as:

$$\tilde{\boldsymbol{z}} = \boldsymbol{J}\boldsymbol{y} \tag{6}$$

where $\tilde{\boldsymbol{z}} = [U_{i,PV}, P_i, Q_i]^{\mathrm{T}} \in \mathbb{R}^{2(N-1)}$ is the auxiliary boundary vector; $\boldsymbol{y} = [U_i, R_{ij}, I_{ij}]^{\mathrm{T}} \in \mathbb{R}^{N-1+2b}$ is the auxiliary state vector; and $\boldsymbol{J} \in \mathbb{R}^{2(N-1)\times(N-1+2b)}$ is a constant matrix, whose elements are determined by the network topology and parameters.

According to (6), the following ELRE can be obtained by:

$$\boldsymbol{y} = \boldsymbol{H}\tilde{\boldsymbol{z}} + \boldsymbol{v} \tag{7}$$

where $\boldsymbol{H} \in \mathbb{R}^{(N-1+2b)\times 2(N-1)}$ is the unknown constant mapping matrix depending on the network topology and network parameters, which is obtained in the offline learning stage; $\boldsymbol{v} \in \mathbb{R}^{N-1+2b}$ is the possible error matrix in historical/simulation data with the expected values $E(\boldsymbol{v}) = 0$.

Multiple historical or simulation PF snapshots are used to learn $\boldsymbol{H}$. Suppose s historical or simulation PF snapshots with the same topology are available, and each snapshot includes the given boundary conditions and the results of the PF calculation. The auxiliary boundary vectors and the auxiliary state vectors of all s historical or simulation PF snapshots are aggregated into:

$$\boldsymbol{Z} = [\tilde{\boldsymbol{z}}_1 \quad \tilde{\boldsymbol{z}}_2 \quad \ldots \quad \tilde{\boldsymbol{z}}_s] \in \mathbb{R}^{2(N-1)\times s} \tag{8}$$

$$\boldsymbol{Y} = [\boldsymbol{y}_1 \quad \boldsymbol{y}_2 \quad \ldots \quad \boldsymbol{y}_s] \in \mathbb{R}^{(N-1+2b)\times s} \tag{9}$$

where $\tilde{\boldsymbol{z}}_i \in \mathbb{R}^{2(N-1)}$ and $\boldsymbol{y}_i \in \mathbb{R}^{N-1+2b} (i = 1, 2, \ldots, s)$ are the auxiliary boundary vector and the auxiliary state vector in the i^{th} historical or simulation PF snapshot, respectively.

According to (7), the ELRE between $\boldsymbol{Y}$ and $\boldsymbol{Z}$ is obtained as:

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{V} \tag{10}$$

where $\boldsymbol{V} \in \mathbb{R}^{(N-1+2b)\times s}$ is the error matrix and $E(\boldsymbol{V}) = 0$.

B. Offline Learning Stage of DDPF

The task of the offline learning stage is to aggregate the historical or simulation PF snapshots with the same topology which will be addressed in Section III-D, and then obtain the mapping matrix $\boldsymbol{H}$ addressed below.

If the row of matrix $\boldsymbol{Z}$ is full rank, the weighted least square (WLS) method can be used to estimate the $\boldsymbol{H}$ matrix directly. However, the matrix $\boldsymbol{Z}$ obtained from historical or simulation PF data may not meet the condition of full row rank, i.e., the collinearity problem may exist in the historical or simulation PF data. To suppress the effect of collinearity of historical or simulation data, the ridge regression (RR) method is used to estimate $\boldsymbol{H}$ corresponding to each topology:

$$\hat{\boldsymbol{H}} = \operatorname{argmin} \sum_{i=1}^{s} (\boldsymbol{y}_i - \boldsymbol{H}\tilde{\boldsymbol{z}}_i)^{\mathrm{T}} (\boldsymbol{y}_i - \boldsymbol{H}\tilde{\boldsymbol{z}}_i) + \lambda \| \boldsymbol{H} \|_F^2 = \boldsymbol{Y}\boldsymbol{Z}^{\mathrm{T}} (\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{T}} + \lambda \boldsymbol{I})^{-1} \tag{11}$$

where $\hat{\boldsymbol{H}}$ is the estimated value of $\boldsymbol{H}$; $\| \cdot \|_F$ is the Frobenius norm; λ is a tuning parameter and $\lambda > 0$; and $\boldsymbol{I}$ is an identity matrix with dimension $2(N-1)$.

Model (11) can be run offline to obtain the estimated mapping matrix $\hat{\boldsymbol{H}}$. The ratio of the number of historical/simulation PF snapshots s to the number of auxiliary state variables $2(N-1)$ affects the calculation accuracy of $\hat{\boldsymbol{H}}$. Obviously, it is preferable to ensure that $s \geq 2(N-1)$ historical/simulation PF snapshots are available.

C. Online Computing Stage of DDPF

The task of the online computing stage is to quickly find the historical or simulation power flow snapshots with the same topology as the current snapshot so as to obtain the corresponding $\hat{\boldsymbol{H}}$ matrix. Then, the auxiliary state vector of the current snapshot can be obtained by:

$$\hat{\boldsymbol{y}}_{new} = \hat{\boldsymbol{H}}\tilde{\boldsymbol{z}}_{new} \tag{12}$$

where $\tilde{\boldsymbol{z}}_{new}$ and $\hat{\boldsymbol{y}}_{new}$ are the auxiliary boundary vector and the estimated auxiliary state vector of the current PF snapshot, respectively.

Note that PF calculation is for the auxiliary state variables, and the estimated values of the original state variables of PF can be obtained. In a word, the main computational burden of the online computing stage is the matrix multiplication, which is less than that of the iterative solution of nonlinear algebraic equations in traditional NRPF or FDPF methods.

D. Topology Processing

Obviously, it is necessary to judge whether different snapshots have the same topology in both the offline learning stage and online computing stage.

According to [4], if $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$ are strongly correlated, it can be considered that $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$ have the same topology. The Pearson correlation coefficient (PCC) is used to measure the correlation. The PCC of $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$ is calculated as:

$$P(\tilde{\boldsymbol{z}}_i,\tilde{\boldsymbol{z}}_j)=\frac{\sum_{k=1}^{m}(\tilde{z}_{i,k}-\bar{\tilde{z}}_i)(\tilde{z}_{j,k}-\bar{\tilde{z}}_j)}{\sqrt{\sum_{k=1}^{m}(\tilde{z}_{i,k}-\bar{\tilde{z}}_i)^2}\sqrt{\sum_{k=1}^{m}(\tilde{z}_{j,k}-\bar{\tilde{z}}_j)^2}} \tag{13}$$

where $\tilde{\boldsymbol{z}}_i\in\mathbb{R}^m$; $\tilde{\boldsymbol{z}}_j\in\mathbb{R}^m$; $m=2(N-1)$; $\tilde{z}_{i,k}$ and $\tilde{z}_{j,k}$ are the k^{th} elements of $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$, respectively; $\bar{\tilde{z}}_i=\sum_{k=1}^{m}\tilde{z}_{i,k}/m$; $\bar{\tilde{z}}_j=\sum_{k=1}^{m}\tilde{z}_{j,k}/m$; and $P(\tilde{\boldsymbol{z}}_i,\tilde{\boldsymbol{z}}_j)$ is the PCC between $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$, and if $P(\tilde{\boldsymbol{z}}_i,\tilde{\boldsymbol{z}}_j)\geq\eta$, then $\tilde{\boldsymbol{z}}_i$ and $\tilde{\boldsymbol{z}}_j$ can be considered to have the same topology, the recommended threshold η is 0.90.

In offline learning stage, the historical/simulation snapshots with the same topology are clustered based on (13). In the online computing stage, the historical/simulation snapshots with the same topology as the current snapshot are also found based on (13), and then the corresponding $\hat{\boldsymbol{H}}$ can be obtained.

Remark: because there is no need for nonlinear iterative computation in the offline learning stage and online computing stage, the proposed DDPF method has no convergence problem.

IV. Case Studies

The performance of the proposed DDPF method is tested and compared with NRPF, FDPF, and two existing DDPF methods in [6], [7] on IEEE transmission and distribution systems on an Intel(R) Core(TM) i5 PC, 2.30 GHz processor with 8 GB RAM. The algorithms are implemented in MATLAB, and NRPF and FDPF are calculated by Matpower (convergence accuracy is 10^{-8}). In all tests, λ in (11) is set to be 10^{-6}.

To verify the correctness of the topology identification method based on PCC, 6 sets of auxiliary boundary vectors are taken from the IEEE 300-bus system, which correspond to three different topologies for the test, i.e., $\tilde{\boldsymbol{z}}_1$ and $\tilde{\boldsymbol{z}}_4$, $\tilde{\boldsymbol{z}}_2$ and $\tilde{\boldsymbol{z}}_5$, and $\tilde{\boldsymbol{z}}_3$ and $\tilde{\boldsymbol{z}}_6$, respectively, having the same topology. The PCC correlation matrix is calculated and shown in Fig. 1, in which it can be observed that the PCC method can accurately identify the same topology. Further, the topological identification results of all other IEEE systems are all correct.

To test the effect of the number of snapshots used in the offline learning stage on the estimation accuracy of the proposed DDPF method, let $R_{atio}=s/[2(N-1)]$. For different ratios (from 0.1 to 1.2), the mean absolute error of voltage magnitude (denoted by $|dV|_m$) and the mean absolute error of phase angle (denoted by $|d\theta|_m$) between the true values and the proposed results of DDPF are given in Figs. 2 and 3. It can be observed that as R_{atio} increases, both $|dV|_m$ and $|d\theta|_m$ gradually decrease.

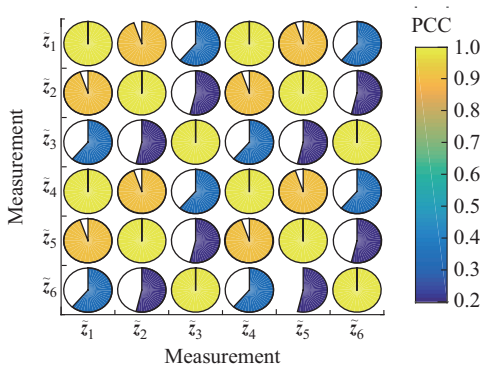

Fig. 1. PCC correlation matrix of 6 sets of auxiliary boundary vectors.

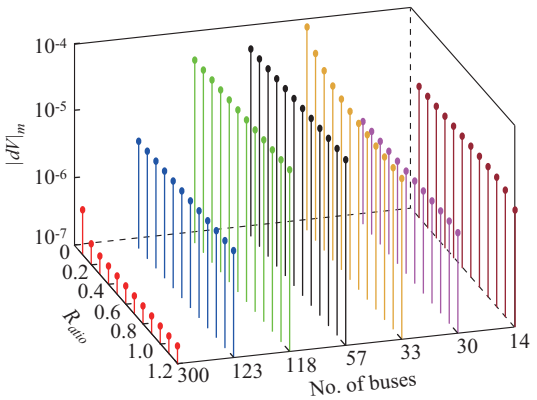

Fig. 2. $|dV|_m$ with R_{atio} for different IEEE systems.

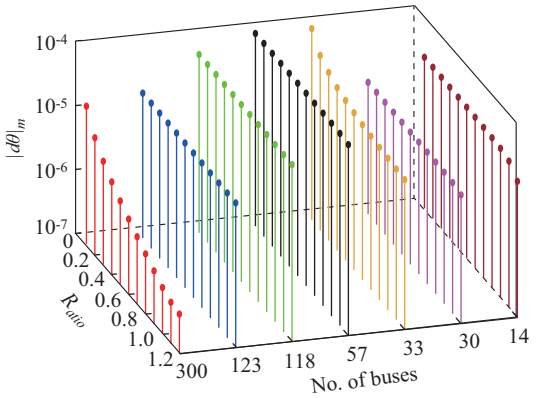

Fig. 3. $|d\theta|_m$ with R_{atio} for different IEEE systems.

When $R_{atio}=1$, both $|dV|_m$ and $|d\theta|_m$ obtained by the proposed DDPF method are smaller than those given in [6] and [7] as shown in Table I, thereby proving that the proposed DDPF method has higher calculation accuracy. Note that even when $R_{atio}=0.2$, both $|dV|_m$ and $|d\theta|_m$ are smaller than 10^{-4}, which indicates that the proposed DDPF method can still work well with relatively small number of historical/simulation PF snapshots.

Suppose the uncertainty of power generations and loads makes the voltage magnitudes vary randomly between 0.95 and 1.05 and the phase angles vary randomly between −15° and 15°. One hundred tests are performed and the state vector of the previous snapshot is used as the initial guess for NRPF and FDPF.

TABLE I
$|dV|_m$ AND $|d\theta|_m$ OBTAINED BY DIFFERENT DDPF METHODS WHEN $R_{atio}=1$

System	DDPF in [6]		DDPF in [7]		Proposed DDPF	
	$\|dV\|_m$	$\|d\theta\|_m$	$\|dV\|_m$	$\|d\theta\|_m$	$\|dV\|_m$	$\|d\theta\|_m$
IEEE 14	7×10^{-5}	4×10^{-5}	1×10^{-5}	1×10^{-3}	8×10^{-6}	2×10^{-5}
IEEE 30	1×10^{-4}	8×10^{-4}	1×10^{-5}	2×10^{-3}	3×10^{-6}	1×10^{-5}
IEEE 33	2×10^{-4}	1×10^{-4}	7×10^{-6}	4×10^{-4}	3×10^{-6}	2×10^{-5}
IEEE 57	6×10^{-4}	2×10^{-4}	2×10^{-4}	3×10^{-2}	6×10^{-5}	9×10^{-5}
IEEE 118	1×10^{-3}	8×10^{-4}	1×10^{-4}	7×10^{-2}	5×10^{-5}	5×10^{-5}
IEEE 123	2×10^{-3}	9×10^{-4}	3×10^{-6}	3×10^{-4}	3×10^{-6}	2×10^{-4}
IEEE 300	2×10^{-3}	4×10^{-4}	2×10^{-4}	2×10^{-4}	2×10^{-7}	5×10^{-7}

The number of times that NRPF and FDPF do not converge in 100 tests (denoted by η) is given in Table II. As can be seen from Table II, NRPF and FDPF have the difficulty in the convergence for a large percentage of the tests when there is an uncertainty in power generations and loads, whereas the proposed DDPF method obtains accurate enough results in all cases just as those in Table I. Also in Table II, the average number of iterations, the average computation time of the DDPF method in [6], and the proposed DDPF method are compared with those of NRPF and FDPF when they converge. Since the online computing stage of DDPF methods do not require iterative calculation, the number of iterations is 0. As can be seen from Table II, among the listed PF methods, the online computing of the proposed DDPF method has much higher computing efficiency than the other methods and is thus more suitable for large-scale networks.

TABLE II
CALCULATION RESULTS OF SEVERAL PF ALGORITHMS UNDER UNCERTAINTY OF GENERATION AND LOAD IN 100 TESTS

System	η	Average number of iterations				Computation time (ms)			
		NRPF	FDPF	DDPF in [6]	Proposed DDPF	NRPF	FDPF	DDPF in [6]	Proposed DDPF
IEEE 14	23	2	6	0	0	11.61	10.52	1×10^3	0.08
IEEE 30	66	3	11	0	0	14.87	13.25	3×10^3	0.09
IEEE 33	78	4	10	0	0	16.24	14.57	3×10^3	0.09
IEEE 57	91	3	7	0	0	20.07	17.91	4×10^3	0.18
IEEE 118	88	3	8	0	0	34.63	28.43	6×10^3	0.41
IEEE 123	84	4	10	0	0	40.12	33.78	6×10^3	0.43
IEEE 300	92	5	9	0	0	76.47	59.39	25×10^3	1.44

Furthermore, in 100 continuous PF sampling snapshots, the topology is assumed to change continuously. For each sampling snapshot, NRPF, FDPF, the DDPF method in [7], and the proposed DDPF method are tested. In all the tests, the state vector of the previous snapshot is used as the initial guess of the current snapshot for NRPF and FDPF. The number of times that NRPF and FDPF do not converge in 100 tests is given in Table III. As can be seen from Table III, NRPF and FDPF might have difficulty in convergence when the topology changes frequently. Note that the DDPF method in [7] cannot be used when the topology changes. Therefore, the average number of iterations and computation time are expressed as ∞ in Table III. Whereas in all tests, the proposed DDPF method correctly finds the historical/simulation snapshots with the same topology as the current snapshot, and then obtains the PF solutions with the same accuracy as that of Table I, the average number of iterations and the computation time of the proposed DDPF method are also given in Table III, thereby proving that the proposed DDPF method can quickly obtain high-precision PF solutions even when the topology changes frequently.

TABLE III
CALCULATION RESULTS OF SEVERAL PF METHODS WHEN TOPOLOGY CHANGES FREQUENTLY IN 100 TESTS

Systems	η	Average number of iterations				Computation time (ms)			
		NRPF	FDPF	DDPF in [7]	Proposed DDPF	NRPF	FDPF	DDPF in [7]	Proposed DDPF
IEEE 14	45	2	6	∞	0	11.28	10.79	∞	0.09
IEEE 30	39	3	10	∞	0	14.19	12.09	∞	0.11
IEEE 33	37	4	10	∞	0	16.93	14.88	∞	0.12
IEEE 57	31	4	8	∞	0	26.67	19.13	∞	0.25
IEEE 118	23	4	8	∞	0	39.89	28.17	∞	0.53
IEEE 123	22	4	10	∞	0	40.57	33.61	∞	0.61
IEEE 300	20	5	9	∞	0	77.61	58.98	∞	1.88

V. CONCLUSION

In the emerging smart grids, traditional MDPF methods may have convergence problems because of the intermittent and uncertain power generations and loads as well as the frequent changes of the topology. At the same time, the accuracy of traditional MDPF methods may also be affected by in-

accurate parameters. In order to solve the above problems, this letter proposes a novel DDPF method including offline learning stage and online computing state which can make full use of historical/simulation PF big data. The proposed DDPF method has the following advantages: ① it has no convergence problem and has very high computational efficiency; ② it can deal with different topological structures and can adapt to the frequent changes of topological structures in smart grids.

Note that there may be gross errors in real-time power flow snapshot caused by accidental fault or malicious attack. In these circumstances, it may be impossible to get accurate state vector by using the proposed DDPF method directly. To this end, we can combine the proposed method with state estimators so as to obtain the estimation value of state vector quickly and accurately. Also, the proposed DDPF method can be extended to the integrated energy system (IES) so as to solve the problem of fast and accurate multi-energy calculation of IES in uncertainty circumstances. These are our future research areas.

REFERENCES

[1] W. F. Tinney and C. E. Hart, "Power flow solution by Newton's method," *IEEE Transactions on Power Apparatus and Systems*, vol. 86, no. 11, pp. 1449-1460, Nov. 1967.
[2] W. F. Tinney and J. W. Walker, "Direct solution of sparse network equations by optimally ordered triangular factorization," *Proceedings of the IEEE*, vol. 55, no. 11, pp. 1801-1809, Nov. 1967.
[3] B. Stott and O. Alsac, "Fast decoupled load flow," *IEEE Transactions on Power Apparatus and Systems*, vol. 91, no. 3, pp. 859-869, May 1974.
[4] Y. Weng, R. Negi, C. Faloutsos *et al.*, "Robust data-driven state estimation for smart grid," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1956-1967, Jul. 2017.
[5] C. Chen, M. Cui, F. Li *et al.*, "Model-free emergency frequency control based on reinforcement learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2336-2346, Apr. 2021.
[6] J. Yu, Y. Weng, and R. Rajagopal, "Mapping rule estimation for power flow analysis in distribution grids," in *Proceedings of 2017 Nath American Power Symposium*, Boston, USA, Sept. 2017, pp. 1-6.
[7] Y. Liu, N. Zhang, Y. Wang *et al*., "Data-driven power flow linearization: a regression approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2569-2580, May 2019.
[8] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458-1459, Aug. 2006.

Yanbo Chen received the B.E. degree in electrical engineering, from Huazhong University of Science and Technology, Wuhan, China, in 2007, the M.S. degree in electrical engineering, from China Electric Power Research Institute, Beijing, China, in 2010, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013. He is currently an Assistant Professor with North China Electric Power University, Beijing, China. His research interests include situation awareness and power system analysis and control.

Chao Wu received the B.S. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2019. He is currently persuing the M.S. degree at North China Electric Power University. His research interests include state estimation and power system analysis and control.

Junjian Qi received the B.E. degree in electrical engineering, from Shandong University, Jinan, China, in 2008, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013. He was a Visiting Scholar with Iowa State University, Ames, USA, in 2012, a Research Associate with the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, USA, from 2013 to 2015, a Postdoctoral Appointee with the Energy Systems Division, Argonne National Laboratory, Lemont, USA, from 2015 to 2017, and an Assistant Professor with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, USA, from 2017 to 2020. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, USA. His research interests include cascading blackouts, microgrid control, cyber-physical systems, and synchrophasors.