Abstract—This paper presents a novel algorithm for recovering missing data of phasor measurement units (PMUs). Due to the low-rank property of PMU data, missing measurement recovery can be formulated as a low-rank matrix-completion problem. Based on maximum-margin matrix factorization, we propose an efficient algorithm based on alternating direction method of multipliers (ADMM) for solving the matrix completion problem. Comparing to existing approaches, the proposed ADMM based algorithm does not need to estimate the rank of the target data matrix and provides better performance in computation complexity. In addition, we consider the case of measurements missing from all PMU channels and provide a strategy of reshaping the matrix which contains the received PMU data for recovery. Numerical results using PMU measurements from IEEE 68-bus power system model illustrate the effectiveness and efficiency of the proposed approaches.

Index Terms—Missing data recovery, ADMM, low-rank matrix completion, phasor measurement units

I. INTRODUCTION

The wide-area measurement system (WAMS) using phasor measurement units (PMUs) has been regarded as one of the key enabling technologies in monitoring, control, and protection of the next-generation power grids [1]. With continuous increase in PMU deployment and the resulting explosion in data volume, the design and deployment of an efficient wide area communication and computing infrastructure, especially from the point of view of resilience against a large number of missing data, is evolving as one of the greatest challenges to the power system and IT communities. With thousands of networked PMUs being scheduled to be installed in the United States by 2020, exchange of synchrophasor data between balancing authorities for any type of wide-area control and stability of the system. Therefore, recovering missing PMU measurements has become a significant and inevitable problem in power systems.

PMU data can be structured as a matrix with each column and row representing the measurements of one channel and sample instant, respectively. Since large amounts of PMU data exhibit heavily correlated property [2]–[4], the matrix is approximately low-rank, and the problem of recovering the missing PMU data can be formulated as a low-rank matrix-completion problem. Studies on matrix completion algorithms are extensive, including ADMiRa [5], singular value projection (SVP) [6], information cascading matrix completion (ICMC) [7], among which nuclear-norm-regularized matrix approximation (NNMRA) [8]–[11] and maximum-margin matrix factorization (MMMF) [12] are widely adapted. Using nuclear-norm-regularized matrix approximation, a singular value thresholding approach was proposed for solving the matrix completion problem [13], [14]. However, the calculation of the singular value decomposition (SVD) in ADM approach increases the computational time and complexity. Based on MMMF, Jain et al. [15] and Hardt [16] proposed alternating least squares (ALS) schemes for solving the matrix-completion problem. Further, softImpute-ALS is provided for reducing the computational complexity [17]. Gao et al. applied the MMMF approach on recovering the missing PMU data [18] firstly. Most of the existing approaches rely on estimation of the rank $r$ of the data matrix, which is typically unavailable and time variant in practice. Inaccurate estimation of $r$ introduces modelling errors in the matrix completion problem. The computational complexity is lower with a smaller $r$. On the other hand $r$ cannot be too small for recovery accuracy. Therefore, design of an adaptive and scalable online algorithm of PMU data recovery is an open challenge.

Motivated by these insights, we develop an algorithm that can recover the missing PMU measurement with low computational complexity and less operating time. The fundamental set-up for this optimization was based on MMMF and alternating direction method of multipliers (ADMM) [19]–[21]. Firstly, the observed PMU data is structured as a matrix $M \in \mathbb{R}^{n_1 \times n_2}$ whose columns and rows represent the measurements from one channel and the same sampling instant, respectively. Then we formulate the data recovery as an optimization problem in which we minimize the rank of the recovery matrix $\hat{X}$ while keeping elements in $X$ the
same as the corresponding ones in \( M \) if they are present. An ADMM algorithm is proposed to solve the optimization problem in an iterative way. In the update equations there is no matrix inverse computation, which immensely reduces the computational complexity. In addition, it is not necessary to estimate the rank of the original data matrix \( X \) without missing elements, which significantly cuts down the influence of the uncertain factor into the performance. Furthermore, we consider the case of missing data from all PMU channels. In this case, all elements in one row of the observed matrix \( M \) are missing. One efficient algorithm is presented to reshape the observed matrix, and the lost data from all the channels can be recovered using ADMM approach. We illustrate the results using simulations of the IEEE 68-bus system model.

II. Problem Formulation

In this section, we present the low-rank property of PMU data and formulate the data recovery as a matrix completion problem.

A. Low-rank property of PMU measurements

Denote \( X \in \mathbb{R}^{n_1 \times n_2} \) as the PMU measurement matrix without data missing. Each column and row correspond to a sequence of measurements of one PMU channel, and the PMU measurements at the same sampling instant, respectively. Due to the noise, all the singular values of \( X \) are larger than zero. An approximating rank approach, referred to Frobenius norm proportion \([22]\), is stated as follows.

\[
\frac{\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}}{\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}} \geq \beta, \tag{1}
\]

where \( \sigma_1 > \sigma_2 > \ldots > \sigma_{n_1} \) are the singular values of the matrix and \( 0 < \beta \leq 1 \), is the proportion factor. \( r \) in (1) denotes the approximate rank of the matrix. Since the PMU measurements of voltage or current phasors or magnitudes from different lines or buses are strongly correlated, the approximate rank of \( X \) is much smaller than \( \min\{n_1, n_2\} \) \([2]–[4]\). Due to the low-rank property of PMU data, missing PMU measurement recovery can be converted into a low-rank matrix completion problem.

B. An ADMM based approach for PMU data recovery

Let \( M \in \mathbb{R}^{n_1 \times n_2} \) and \( \hat{X} \in \mathbb{R}^{n_1 \times n_2} \) denote the observed PMU measurements with missing data and the recovered matrix, respectively. Since \( \hat{X} \) should be a low-rank matrix, the matrix completion problem is formulated as follows:

\[
\begin{align*}
\min_{X \in \mathbb{R}^{n_1 \times n_2}} & \quad \text{rank}(X) \\
\text{subject to} & \quad (X - M) \odot I_s = 0,
\end{align*} \tag{2}
\]

where \( \odot \) denotes the Hadamard product, \( I_s \) is the structural identity with its \( i_j^{th} \) entry defined as

\[
[I_s]_{ij} = \begin{cases} 
1, & \text{if } [M]_{ij} \text{ is observed data;} \\
0, & \text{if } [M]_{ij} \text{ is missing data.}
\end{cases}
\tag{3}
\]

Unfortunately, (2) is NP hard to solve, and can be relaxed to a tractable optimization problem \([23]\):

\[
\begin{align*}
\min_{X \in \mathbb{R}^{n_1 \times n_2}} & \quad \|X\|_* \\
\text{subject to} & \quad (X - M) \odot I_s = 0,
\end{align*} \tag{4}
\]

where the nuclear norm \( \|X\|_* \) is the sum of the singular values of \( X \).

Using MMMF to further change the optimization problem \([4]\), let \( \hat{X} = A^T B \), in which \( A \in \mathbb{R}^{n_2 \times n_1} \) and \( B \in \mathbb{R}^{n_2 \times n_2} \). Without loss of generality, we assume \( n_1 > n_2 \). Since \( \|X\|_* \) is equivalent to \( \min_{A,B} \frac{1}{2} (||A||_F^2 + ||B||_F^2) \) with Frobenius norm \( ||.||_F \) \([12]\), the optimization function is equivalent to

\[
\begin{align*}
\min_{A,B} & \quad \frac{1}{2} (||A||_F^2 + ||B||_F^2) \\
\text{subject to} & \quad (A^T B - M) \odot I_s = 0.
\end{align*} \tag{5}
\]

In the previous work \([13]–[16]\), people estimated the rank \( r \) of \( \hat{X} \), set \( A \in \mathbb{R}^{n_2 \times n_1} \) and \( B \in \mathbb{R}^{n_2 \times n_2} \), and applied ALS to solve (5). The computational complexity is \( O((n_1 + n_2)^r) \). If \( r = \min\{n_1, n_2\} \), the computational complexity is \( O((n_1 + n_2)(\min\{n_1, n_2\} )^3) \), which is a biquadratic function of \( \min\{n_1, n_2\} \). With smaller \( r \) the computational complexity is reduced. However, the value of \( r \) cannot be too small to guarantee the recovery accuracy. For reducing the influence of the uncertain factor into the performance, we set the sizes of matrices \( A \in \mathbb{R}^{n_2 \times n_1} \) and \( B \in \mathbb{R}^{n_2 \times n_2} \) only depend on the size of observed matrix \( M \). In addition, we apply the ADMM method to solve (5) in an iterative way using the Lagrangian multiplier approach.

The augmented Lagrangian for (5) can be formulated as

\[
L = \frac{1}{2} (||A||_F^2 + ||B||_F^2) + \text{trace}(w^T (A^T B - M) \odot I_s) + \frac{\rho}{2} ||(A^T B - M) \odot I_s||_F^2, \tag{6}
\]

where \( A \) and \( B \) are the matrices of the primal variables, \( w \) is the matrix of the dual variables or the Lagrange multipliers associated with (5), and \( \rho > 0 \) denotes a penalty weight.

After some algebraic, the augmented Lagrangian can be rewritten as

\[
L = \frac{1}{2} (||A||_F^2 + ||B||_F^2) + \text{trace}((w \odot I_s)^T (A^T B - M)) + \frac{\rho}{2} \text{trace}(((A^T B - M) \odot I_s)^T (A^T B - M)). \tag{7}
\]

The gradients of the augmented Lagrangian \( L \) in (7) with respect to \( A \) and \( B \) are respectively given by

\[
\frac{\partial L}{\partial A} = A + B(w \odot I_s)^T + \rho B((A^T B - M) \odot I_s)^T, \\
\frac{\partial L}{\partial B} = B + A(w \odot I_s) + \rho A(A^T B - M) \odot I_s. \tag{8}
\]

Given the derivation, the ADMM algorithm for solving the optimal problem (5) is illustrated in Algorithm 1.

The updates in Algorithm 1 requires no matrix inverse, and the computational complexity is \( O(n_1 n_2 \min\{n_1, n_2\}) \), which is a quadratic function of \( \min\{n_1, n_2\} \). In addition, it is not necessary to estimate the rank of matrix \( \hat{X} \), which reduces the
The influence of uncertain factor into the performance. The convergent rate of using ADMM depends on the penalty weight \( \rho \). Due to less computational complexity, \( \rho \) can be chosen as a smaller number for improving the recovery accuracy.

C. Special case: missing data from all the channels

The power system often suffers natural and artificial disturbances during operation. It is possible that the data from all the channels are missing simultaneously under communication failure. In this case, no existing algorithms can recover the missing data. For solving this problem, the observed matrix \( M \) has to be reshaped to avoid some of its rows missing. Our goal is that the proportion of missing elements in one row of the reshaped observed matrix \( \tilde{M} \) is as small as possible. Meanwhile the corresponding reshaped recovery matrix \( \tilde{X} \) is still low-rank.

We provide an alternative method, called cut-column reshaping method (CCRM), for reshaping the observed matrix. Using CCRM each column with \( n_1 \) length is separated into \( n^* \) shorter columns with a length of \( \frac{n_1}{n^*} \). Thus, the \( n_1 \)-by-\( n_2 \) matrix is reshaped to a \( \frac{n_1}{n^*} \)-by-\( n_2 n^* \) matrix, and the original column correlation is held. The length of the new column should be larger than the row length of the original matrix, \( i.e., \frac{n_1}{n^*} > n_2 \). \( n^* \) also satisfies that \( \frac{n_1}{n^*+1} < n_2 \), where \( \lfloor x \rfloor \) denotes the smallest integer number which is larger than \( x \). Thus the numbers of rows and columns of reshaped matrix are both larger than \( n_2 \). Due to the size, the rank of \( M \) is no more than \( \min\{n_1, n_2\} \). Using CCRM the rank of reshaped matrix \( \tilde{M} \) will not be reduced by the new size. In addition, with holding the column correlation, CCRM minimizes the proportion of zero elements in one row of reshaped matrix.

Consider a simple example to illustrate the reshaping method. A 6-by-2 matrix \( M \) can be expressed as:

\[
M = \begin{bmatrix}
m_1 & m_2 \\ m_{11} & m_{21} & m_{41} & m_{41} & m_{61} \\ m_{12} & m_{22} & m_{32} & m_{42} & m_{62}
\end{bmatrix}^T
\]

whose fifth row is missing. Using CCRM with \( n^* = 3 \) and matrix \( M \) is reshaped into a 2-by-6 matrix:

\[
\tilde{M} = \begin{bmatrix}
m_1 & \tilde{m}_2 & \tilde{m}_3 & \tilde{m}_4 & \tilde{m}_5 & \tilde{m}_6 \\ m_{11} & m_{41} & m_{12} & m_{32} & m_{61} & m_{62}
\end{bmatrix}
\]

Now for each column and row, not all measurements are missing. If \( m_1 \) and \( m_2 \) are strongly correlated, \( m_1 \) and \( \tilde{m}_4 \), \( \tilde{m}_2 \) and \( \tilde{m}_5 \), and \( \tilde{m}_3 \) and \( \tilde{m}_6 \) are strongly correlated in pairs.

The ranks of matrix \( M \) and \( \tilde{M} \) are both no more than 2. The proportion of missing elements to the first row is \( \frac{2}{3} \); while it is 1 to the fifth row of \( \tilde{M} \). CCRM is illustrated in Algorithm 2. The missing PMU measurements from all the channels can be recovered using ADMM in Algorithm 1 after reshaped matrix \( M \) using CCRM in Algorithm 2.

### III. Simulation results

The IEEE 68-bus system is used to carry out the simulation to verify the proposals. We build up a PMU measurement matrix whose column and row corresponding to a sequence voltage phasors on 86 lines and the sampling instant, respectively. The simulated measurements are obtained using the power systems toolbox (PST) nonlinear dynamics simulation routine s_simu and the data file data16m.m [24]. A threephase fault is imposed at the line connecting buses 1 and 2. The fault starts at \( t = 0.1s \), and clears at bus 1 at \( t = 0.15s \) and at bus 2 at \( t = 0.20s \). For approaching to the true measurements, we add white Gaussian noise \( \mathcal{N}(0, 0.001) \) into the PMU data. The measurements are observed during 60s and there are 30 samples in one second. The 1800-by-86 matrix \( X \) is with no missing measurements and its approximate rank is 1 with \( \beta = 0.995 \) in (1). To test the recovery accuracy of the presented ADMM algorithm, some observed data in \( X \) is set to be lost. Since the PMU data are missing arbitrary and unpredictable, in this paper we consider two cases of missing data: (1) Missing data randomly. The delivery of PMU measurements from multiple remote locations of power grids to monitoring centers can result in the random unavailability of PMU measurements; (2) Missing data in all channels simultaneously. The transform link malfunctions may result in data missing in all channels. We choose the penalty weight \( \rho = 0.00075 \) using ADMM, and the dual parameter \( \lambda = 1.5 \) and the estimated rank of filled completion matrix \( r = 20 \) using ALS for comparison. In the paper, the computational time is obtained by operating Matlab programming.

**Case 1: Missing data randomly**

In this case, we assume an independent and identical distribution (i.i.d) of the missing rate. For each data point, with a probability of 0.2, the measurement is missing and set to zero in \( M \) artificially. Notice that it is different from the data which is equal to zero. If the actual data is zero, the
corresponding element in $I_s$ is equal to 1. While if the data is missing, the corresponding element in $I_s$ is equal to 0.

Fig. 1 shows the trajectories of the first elements of the filled completion matrix $\hat{X}$ using ADMM and ALS, respectively. The convergence speed using ALS outperforms to the one using ADMM. The mean absolute errors (MAEs) using ADMM and ALS are 0.0079 and 0.0075, respectively.

Fig. 2 shows the computational time using ADMM and ALS, respectively. Though the convergent rate of ADMM is 100 iterations which is larger than using ALS, the total computational time is around 1 s. While if we use ALS to recover the missing data, the computational time is more than 7s in 50 iterations.

Fig. 3 shows the statistic, maximum, and minimum values of MAEs using ADMM and ALS with different observed data probabilities, respectively. The statistic, maximum, and minimum values of MAEs are obtained by Monte Carlo method with 500 independent times. With larger probability of observed measurements, MAE becomes smaller. The difference between the maximum and minimum values of MAEs using ADMM is larger than the one using ALS.

Case 2: Missing data in all channels

In this case, one row of data in matrix $M$ is lost. The 1800-by-86 matrix $M$ which contains voltage phasor measurements can be treated as 1800 sub-matrices with a size of 1-by-86. The observed data probability denotes the proportion of the observed sub-matrices to the total ones. For recovering the missing data in one row, firstly we reshape the observed matrix using CCRM. The original 1800-by-86 matrix $M$ is reshaped to a 90-by-1720 matrix $\tilde{M}$ with $n^* = 20$. With $\beta = 0.995$ in (1), the approximate rank of the corresponding reshaped observed matrix $\tilde{X}$ is 1.

Fig. 4 shows the trajectories of the first elements of the filled completion matrix $\hat{X}$ with 80% observed data probability using ADMM and ALS, respectively. Similar to Case 1, the convergent speed using ALS is faster than the one using ADMM. The MAEs using ADMM and ALS are 0.0263 and 0.0264, respectively.

Fig. 5 shows the computational time using ADMM and ALS, respectively. Though the convergence rate of ADMM is slower than using ALS, the total computational time is much less.
using neither ADMM nor ALS, and the MAEs with different X recovery accuracy cannot be ignored. If the observed matrix mum singular value becomes larger, whose influence into the performance. We
ADMM algorithm. In addition, the ADMM algorithm avoids noisy measurements from the IEEE 68-bus power system model. Compared with the ALS algorithm, the computational model. Compared with the ALS algorithm, the computational
0.027
0.028
0.029
0.03
0.5
0.6
0.7
0.8
0.9
Average observed probability
MAE
Fig. 5. Case 2: computational time using ADMM and ALS, respectively.

shows the statistic, maximum, and minimum values of MAEs using ADMM and ALS with different observed data probabilities, respectively. The statistic values of MAEs using ADMM and ALS are still close. Compared with Case 1, the MAEs using both ADMM and ALS are larger. Though approximate rank of reshaped matrix $\tilde{X}$ is still 1, the minimum singular value becomes larger, whose influence into the recovery accuracy cannot be ignored. If the observed matrix $X$ is not reshaped, the missing row cannot be recovered using neither ADMM nor ALS, and the MAEs with different observed data probabilities are all around 0.278.

IV. CONCLUSION

In this paper, we presented ADMM algorithm for missing PMU measurement recovery. We illustrated our results with noisy measurements from the IEEE 68-bus power system model. Compared with the ALS algorithm, the computational complexity and operating time are much smaller using the ADMM algorithm. In addition, the ADMM algorithm avoids to estimate the rank of filled completion matrix, which reduces the influence of the uncertain factor into the performance. We also consider the case of missing data in all the channels simultaneously and provide one approach to reshape the observed matrix for the recovery. Our future work in this area will include recovering continuous several rows of the observed matrix with all missing elements and testing the proposal using actual PMU data.

REFERENCES

[1] A. G. Phadke and J. S. Thorp, “Synchronized Phasor Measurements and Their Applications,” New York, Springer, 2008.
[2] Y. Chen, L. Xie, and P. Kumar, “Dimensionality Reduction and Early Event Detection Using Online Synchrophasor Data,” in Proc. IEEE Power and Energy Society General Meeting, pp. 1-5, 2013.
[3] N. Dahal, R. L. King, and V. Madani, “Online Dimension Reduction of Synchrophasor Data,” in Proc. IEEE PES Transmission and Distribution Conf. Expo. (T&D), pp. 1-7, 2012.
[4] P. Gao, M. Wang, S. Ghiocel, and J. H. Chow, “Modeless Reconstruction of Missing Synchrophasor Measurements,” in Proc. IEEE PES General Meeting, pp. 1-5, 2014.
[5] K. Lee and Y. Bresler, “ADMPA: Atomic Decomposition for Minimum Rank Approximation,” IEEE Trans. Inf. Theory, vol. 56, no. 9, pp. 4402-4416, 2010.
[6] P. Jain, R. Meka, and I. S. Dhillon, “Guaranteed Rank Minimization via Singular Value Projection,” in Adv. Neural Inf. Process. Syst., pp. 937-945, 2010.
[7] R. Meka, P. Jain, and I. S. Dhillon, “Matrix Completion from Power-Law Distributed Samples,” in Adv. Neural Inf. Process. Syst., pp. 1258-1266, 2009.
[8] E. J. Candes and T. Tao, “The Power of Convex Relaxation: Near-Optimal Matrix Completion,” IEEE Trans. on Information Theory, vol. 56, no. 5, pp. 2053-2080, 2010.
[9] R. Mazumder, H. Trevor, and T. Robert, “Spectral Regularization Algorithms for Learning Large Incomplete Matrices,” Journal of machine learning research, vol. 11, pp. 2287-2322, 2010.
[10] J. Cai, E. J. Candes, and Z. Shen, “A Singular Value Thresholding Algorithm for Matrix Completion,” SIAM Journal on Optimization, vol. 20, no. 4, pp. 1956-1982, 2010.
[11] S. Ma, D. Goldfarb, and L. Chen, “Fixed Point and Bregman Iterative Methods for Matrix Rank Minimization,” Mathematical Programming, vol. 128., no. 1, pp. 321-353, 2011.
[12] J. D. Rennie and S. Nathan, “Fast Maximum Margin Matrix Factorization for Collaborative Prediction,” in Proc. the 22nd international conference on Machine learning(ACM), pp. 713-719, 2005.
[13] C. Chen, B. He, and X. Yuan, “Matrix Completion via an Alternating Direction Method,” JMA Journal of Numerical Analysis, vol. 32, no. 1, pp. 227-245, 2012.
[14] F. Xu, and P. Pan, “A New Algorithm for Positive Semidefinite Matrix Completion,” Journal of Applied Mathematics, vol. 3, pp. 1-5, 2016.
[15] P. Jain, N. Praneeth, and S. Sujay, “Low-Rank Matrix Completion Using Alternating Minimization.” in Proc. the forty-fifth annual ACM symposium on Theory of computing (ACM), 2013.
[16] M. Hardt,“Understanding Alternating Minimization for Matrix Completion,” Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium, 2014.
[17] T. Hastie, R. Mazumder, J. D. Lee, and R. Zadeh, “Matrix Completion and Low-Rank SVD via Fast Alternating Least Squares,” J. Mach. Learn. Res., vol. 16, no. 1, pp. 3367-3402, 2015.
[18] P. Gao, M. Wang, S. G. Ghiocel, J. H. Chow, B. Fardanesh, and G. Stefopoulos, “Missing Data Recovery by Exploiting Low-Dimensionality in Power System Synchrophasor Measurements,” IEEE Trans. on Power Systems, vol. 31, no. 2, pp. 1006-1013, 2016.
[19] R. Glowinski and A. Marroco, “Sur l’approximation, par lments finis d’ordre un, et la resolution, par balanization-dualit d’une classe de problmes de Dirichlet non lineaires.” ESAIM: Mathematical Modelling and Numerical Analysis - Modlisation Mathmatique et Analyse Numrique 9.R2 pp. 41-76, 1975.
[20] D. Gabay and B. Mercier, “A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation”. Computers & Mathematics with Applications, vol. 2, no. 1, pp. 17-40, 1976.
[21] R. Glowinski and L. T. Patrick, “Augmented Lagrangian and operator-splitting methods in nonlinear mechanics,” Society for Industrial and Applied Mathematics, 1989.
[22] X. Zhang, “Matrix Analysis and Applications,” Tsinghua and Springer Publishing house, Beijing, pp. 71-100, 2004.
[23] M. Fazel, “Matrix Rank Minimization with Applications,” Ph. D. dissertation, Stanford Univ., Stanford, CA, 2002.
[24] J. H. Chow and K. W. Cheung, “A Toolbox for Power System Dynamics and Control Engineering Education and Research,” IEEE Transactions on Power Systems, vol. 7, no. 4, pp. 1559-1564, 1992.