An ADMM-Net for Data Recovery in Wireless Sensor Networks

Liu Yang[1], Yonina C. Eldar[2], Haifeng Wang[1], Kai Kang[3], and Hua Qian[3]
[1] Key Laboratory of Wireless Sensor Network & Communication, SIMIT, CAS
[2] Department of Mathematics and Computer Science, Weizmann Institute of Science
[3] Shanghai Advanced Research Institute, CAS

***Abstract*—Data collection plays an important role in wireless sensor networks. Recovery of spatio-temporal data from incomplete sensing data is vital to the network lifetime. Many works have utilized the spatial and temporal correlations to achieve satisfactory data recovery results. However, these methods introduce large computational overhead at the fusion center. In this paper, we develop an ADMM-Net framework for correlated spatio-temporal data recovery. Both the spatial correlation and temporal correlation of sensing data are considered in a convex optimization problem, which is solved by the alternating direction method of multipliers (ADMM) algorithm. We then unfold the ADMM algorithm into a fixed-length neural network that reduces the iterations dramatically and does not require additional location information of nodes. Experimental results on a real-world dataset demonstrate that the proposed method can achieve faster convergence speed than the baseline ADMM algorithm with slight accuracy loss.**

***Index Terms*—Data recovery, ADMM, unfolding, wireless sensor networks.**

I. Introduction

Data collection is a crucial application of wireless sensor networks (WSNs) [1]. Spatio-temporal signals are time series collected over a certain spatial range [2]. In many cases, the collected spatio-temporal data is incomplete because of the limited power supply or malfunctions [3]. Hence, recovery of spatio-temporal signals from incomplete data is a critical issue in WSNs.

Many studies approximate spatio-temporal signals with low-rank matrices. These approaches claim that complete data can be recovered from a subset given the low-rank nature. In [4], a spatio-temporal compressive data gathering scheme (STCDG) based on matrix completion was presented, which made use of both the low-rank and short-term stability features of the data. In [5], a data collection method was proposed based on low-rank matrix approximation (LRMA), where both the temporal consistency and the spatial correlation were simultaneously integrated. Low rank and differential smoothness based recovery (LRDS) was proposed in [2], which introduced a differential smooth prior of time-varying graph signals to the field of spatio-temporal signal analysis. The above algorithms reduce the communication cost and the sensing cost at the expense of heavier computational burden at the fusion center (FC).

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61671436) and the Science and Technology Commission Foundation of Shanghai (Grant No. 18511103502).

Unfolding was first suggested in [6] to accelerate the convergence of conventional iterative algorithms. The main idea of unfolding is to unroll numerical algorithm into a network architecture based on the resulted iterations. Normally, an iterative algorithm can be considered as a recurrent neural network, in which the l-th iteration is regarded as the l-th layer. Prior works have unrolled the proximal gradient method [7], the iterative shrinkage/thresholding algorithm [8], the primal-dual algorithm [9], and any others [10]. Various fields have benefited from unfolding, such as ultrasound imaging [11], image deblurring [12], and image compressive sensing [13].

In this paper, we develop an ADMM-Net framework for correlated spatio-temporal data recovery in WSNs. Both the temporal consistency and the spatial correlation of sensing data are considered. The alternating direction method of multipliers (ADMM) algorithm is adopted to solve the proposed convex optimization problem. We further unroll and generalize the ADMM algorithm into a fixed-length neural network without known the location of nodes. The proposed ADMM-Net framework is evaluated on a real dataset of a sensor network. The results verify that the proposed method can speed up the convergence process with slight loss of recovery accuracy.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and formulate the data recovery in WSNs as a convex optimization problem. The proposed ADMM-Net is described in Section III. Recovery performance is tested by experimenting on a real-world dataset in Section IV. Finally, Section V concludes the paper.

II. System Model and Problem Formulation

We consider a WSN consisting of M sensor nodes and N rounds of data, which are collected into a measurement matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N] = [\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^M]^T \in \mathbb{R}^{M \times N}$. As illustrated in Fig. 1, due to sensor malfunction or conservation of resources, only a subset of sensor nodes are randomly selected for data sensing and transmission at each time slot [3], so that the measurement matrix is generally incomplete. The sampled signal can be modeled as [2]:

$$\mathbf{Y} = \mathbf{Q} \circ (\mathbf{X} + \mathbf{N}), \tag{1}$$

where $\mathbf{Q} \in \{0, 1\}^{M \times N}$ is the sampling matrix, operation "$\circ$" is the Hadamard (element-wise) product, and $\mathbf{N}$ denotes zero-

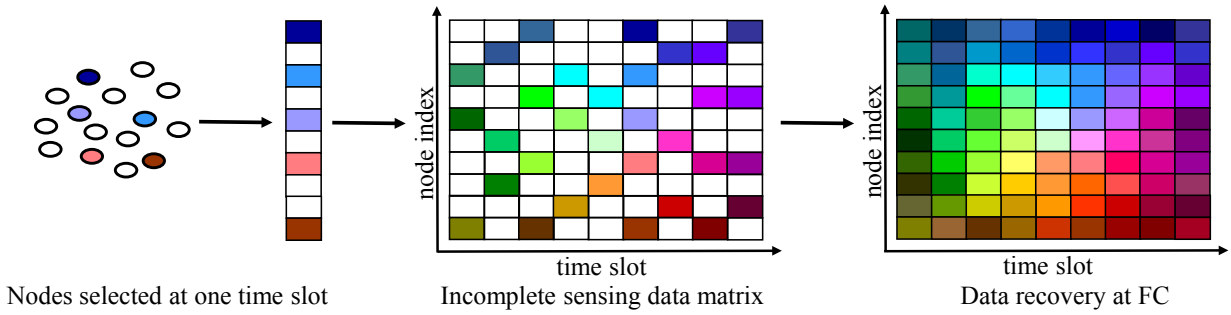

Fig. 1. Correlated spatio-temporal data recovery.

mean Gaussian noise with variance σ^2. The sampling matrix $\mathbf{Q}$ is defined by

$$\mathbf{Q}(i,j) = \begin{cases} 1, & \text{if } i\text{-th node works at } j\text{-th time slot,} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Without loss of generality, the physical quantities describing natural phenomena are locally consistent. In addition, sensor nodes are often deployed redundantly to obtain complete environmental monitoring data. Therefore, sensing data often has high spatial and temporal correlations, which results in a low-rank property of the sensing data [14]. It is known from the low-rank matrix completion theory [15] that it is possible to recover a low-rank matrix from a subset of its entries. Thus, we can attempt to recover $\mathbf{X}$ by solving the following optimization problem [5]:

$$\min_{\mathbf{X}} \ \text{rank}(\mathbf{X}), \quad \text{s.t. } \|\mathbf{Q} \circ \mathbf{X} - \mathbf{Y}\|_F^2 \leq \epsilon, \tag{3}$$

where $\text{rank}(\cdot)$ denotes the rank of a matrix, $\|\cdot\|_F$ represents the Frobenius norm, and ϵ is the error tolerance. However, further intrinsic properties of the measurements are not directly considered in problem (3) beyond the low-rank.

The sequential measurements of a node typically are smooth except for noise and outliers. To model this effect, we introduce the temporal differential operator [2]:

$$\mathbf{D} = \begin{bmatrix} -1 & & \\ 1 & \ddots & \\ & \ddots & -1 \\ & & 1 \end{bmatrix}_{N \times (N-1)}. \tag{4}$$

The temporal difference of $\mathbf{X}$ is then given by $\mathbf{XD} = [\mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_2, \cdots, \mathbf{x}_N - \mathbf{x}_{N-1}]$. Requiring this difference to be small will enforce consistency and smoothness of adjacent columns of the recovered $\mathbf{X}$. In addition, often the measurements of a node can be represented as a linear combination of measurements of its neighbors [14]. Thus, for each $\mathbf{x}^i$, there exists K most correlated rows $\mathbf{x}^{i_k}$, $k = 1, \cdots, K$, such that $\mathbf{x}^i = \sum_{k=1}^{K} \omega_{i_k} \mathbf{x}^{i_k}$, where ω_{i_k} is the spatial correlation weight of $\mathbf{x}^{i_k}$. Denote by $\mathbf{S} \in \mathbb{R}^{M \times M}$ the matrix representing the relation among measurements of different nodes [5]:

$$\mathbf{S}(i,j) = \begin{cases} 1, & \text{if } i = j, \\ -\omega_{i_k}, & \text{if } j = i_k, k = 1, \cdots, K, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The spatial difference of $\mathbf{X}$ is then $\mathbf{SX}$. Requiring it to be small enforces the spatial correlation of the recovered $\mathbf{X}$.

Using these observations, the correlated spatio-temporal data recovery in WSNs is formulated as:

$$\begin{aligned} \min_{\mathbf{X}} \ & \frac{1}{2} \|\mathbf{Q} \circ \mathbf{X} - \mathbf{Y}\|_F^2 + \frac{\lambda_1}{2} \|\mathbf{XD}\|_F^2 \\ & + \frac{\lambda_2}{2} \|\mathbf{SX}\|_F^2 + \mu \|\mathbf{X}\|_*, \end{aligned} \tag{6}$$

where λ_1, λ_2, and μ denote the regularization parameters, $\text{rank}(\mathbf{X})$ is replaced by the nuclear norm $\|\mathbf{X}\|_*$, defined as the sum of the singular values of $\mathbf{X}$, which still promotes the low-rank [16], but in a convex form.

III. Proposed ADMM-Net Framework

A. Traditional ADMM Algorithm

The optimization problem in (6) can be solved efficiently using the ADMM algorithm [17]. By introducing an auxiliary variable $\mathbf{Z}$, the optimization problem (6) is equivalent to

$$\begin{aligned} & \min_{\mathbf{X},\mathbf{Z}} f(\mathbf{X}) + r(\mathbf{Z}), \\ & \text{s.t. } \mathbf{X} = \mathbf{Z}, \end{aligned} \tag{7}$$

where

$$\begin{aligned} f(\mathbf{X}) &= \frac{1}{2}||\mathbf{Q} \circ \mathbf{X} - \mathbf{Y}||_F^2 + \frac{\lambda_1}{2}||\mathbf{XD}||_F^2 + \frac{\lambda_2}{2}||\mathbf{SX}||_F^2, \\ r(\mathbf{Z}) &= \mu||\mathbf{Z}||_*, \end{aligned}$$

in which $f(\cdot)$ is differentiable and $r(\cdot)$ reveals the low-rank property of the sensing data.

The augmented Lagrangian function of (7) is given by

$$\begin{aligned} & \mathcal{L}_\rho(\mathbf{X}, \mathbf{Z}, \mathbf{M}) \\ & = f(\mathbf{X}) + r(\mathbf{Z}) + \langle \mathbf{M}, \mathbf{X} - \mathbf{Z} \rangle + \frac{\rho}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2, \end{aligned} \tag{8}$$

where $\mathbf{M}$ is the Lagrange multiplier, ρ is the penalty parameter, and $\langle \cdot, \cdot \rangle$ denotes the inner product of matrices.

For simplicity, let $\mathbf{P} = \mathbf{M}/\rho$. We perform alternating minimizations and obtain the following iteration schemes:

$$\begin{aligned} & \text{vec}\left(\mathbf{X}^{l+1}\right) \\ & = (\tilde{\mathbf{Q}} + \lambda_1 \tilde{\mathbf{D}} + \lambda_2 \tilde{\mathbf{S}} + \rho \tilde{\mathbf{I}})^{-1} \text{vec}[\rho(\mathbf{Z}^l - \mathbf{P}^l) + \mathbf{Y}], \end{aligned} \tag{9}$$

where $\text{vec}(\cdot)$ denotes the vectorization operator stacking the columns of a matrix into a long vector. Notation $\tilde{\mathbf{I}}$ represents

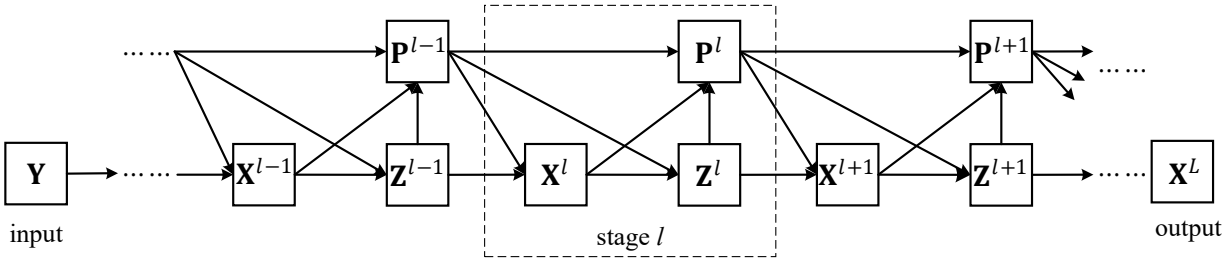

Fig. 2. Data flow graph of the ADMM algorithm.

the MN dimensional identity matrix, $\tilde{\mathbf{Q}} = \text{diag}(\text{vec}(\mathbf{Q}))$, $\tilde{\mathbf{D}} = (\mathbf{D}\mathbf{D}^T) \otimes \mathbf{I}_M$, and $\tilde{\mathbf{S}} = \mathbf{I}_N \otimes (\mathbf{S}^T\mathbf{S})$, where "$\otimes$" denotes the Kronecker product.

The updates of $\mathbf{Z}$ and $\mathbf{P}$ are given by

$$\mathbf{Z}^{l+1} = \text{SVT}_\tau\left(\mathbf{X}^{l+1} + \mathbf{P}^l\right), \text{ where } \tau = \mu/\rho, \tag{10}$$

$$\mathbf{P}^{l+1} = \mathbf{P}^l + \eta\left(\mathbf{X}^{l+1} - \mathbf{Z}^{l+1}\right), \tag{11}$$

where η is the step size, SVT denotes the singular value thresholding operator [2]: $\text{SVT}_\tau(\mathbf{X}) = \mathbf{U}\mathbf{\Lambda}_\tau(\mathbf{\Sigma})\mathbf{V}^T$, where the singular value decomposition of $\mathbf{X}$ is given by $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, and $\mathbf{\Lambda}_\tau(x) = \text{sign}(x)\max(|x| - \tau, 0)$ is the soft thresholding operator defined for $\tau \in \mathbb{R}^+$.

B. ADMM-Net Framework

Although ADMM is a powerful optimization framework, it usually converges slowly with high accuracy requirement, even for simple examples [18]. In addition, the convergence of ADMM relies on the selections of $\{\lambda_1, \lambda_2, \rho, \tau, \eta\}$. To accelerate the convergence speed, in this subsection, we develop an ADMM-Net framework based on the ADMM algorithm to learn these parameters.

We first map the ADMM iterative procedures to a data flow graph [13]. As shown in Fig. 2, different nodes in the graph correspond to different operations in ADMM, the directed edges correspond to the data flows between operations, and the l-th stage of the data flow graph corresponds to the l-th iteration of ADMM. The entire data flow graph is a multiple repetition of the above stages corresponding to successive iterations in ADMM.

The proposed ADMM-Net is defined over the data flow graph. It keeps the graph structure and generalizes the three types of operations as network layers to have learnable parameters. These operations are generalized as:
(1) Reconstruction layer $\mathbf{X}^{l+1}$:

$$\begin{aligned}&\text{vec}\left(\mathbf{X}^{l+1}\right) \\ &= (\tilde{\mathbf{Q}} + \lambda_1^l\tilde{\mathbf{D}} + \lambda_2^l\tilde{\mathbf{S}}^l + \rho^l\tilde{\mathbf{I}})^{-1}\text{vec}[\rho^l(\mathbf{Z}^l - \mathbf{P}^l) + \mathbf{Y}],\end{aligned} \tag{12}$$

where $\tilde{\mathbf{S}}^l = \mathbf{I}_N \otimes [(\mathbf{S}^l)^T\mathbf{S}^l]$. Most methods to construct $\mathbf{S}$ are based on the spatial deployment of sensor nodes or complete measurement matrix, which need additional information or computation. Thus, except for λ_1^l, λ_2^l, and ρ^l, we also learn $\mathbf{S}^l$ from the training data without known the location of nodes.

(2) Nonlinear transform layer $\mathbf{Z}^{l+1}$:

$$\mathbf{Z}^{l+1} = \text{SVT}_{\tau^l}\left(\mathbf{X}^{l+1} + \mathbf{P}^l\right). \tag{13}$$

Given the l-th stage, the thresholding value for SVT operation is given by $\tau^l = \sigma(\nu^l) \cdot \gamma \cdot \max(\Sigma^l)$, where $\sigma(x) = 1/(1 + \exp(-x))$ is the sigmoid function, $\max(\Sigma^l)$ is the maximum singular value of $\mathbf{X}^{l+1} + \mathbf{P}^l$, γ is a fixed scalar (we chose $\gamma = 0.4$ in this paper), and ν^l is learned in ADMM-Net [8].
(3) Multiplier update layer $\mathbf{P}^{l+1}$:

$$\mathbf{P}^{l+1} = \mathbf{P}^l + \eta^l\left(\mathbf{X}^{l+1} - \mathbf{Z}^{l+1}\right), \tag{14}$$

where η^l is learned in ADMM-Net.

In summary, the learnable parameters in ADMM-Net are $\left\{\mathbf{S}^l, \lambda_1^l, \lambda_2^l, \rho^l, \nu^l, \eta^l\right\}_{l=1}^L$. These parameters are learned independently for each layer.

We choose the normalized mean square error (NMSE) as the loss function in network training. Given pairs of training data, the loss between the network output and the ground-truth is defined as [19]:

$$\text{NMSE} = \frac{1}{|\Gamma|}\sum\frac{\left\|\mathbf{X} - \mathbf{X}^L\right\|_F^2}{\|\mathbf{X}\|_F^2}, \tag{15}$$

where $\mathbf{X}^L$ is the network output, $\mathbf{X}$ is regarded as the ground-truth, and the set Γ represents the pairs of training data.

To obtain the optimal parameters, the backpropagation strategy is adopted in ADMM-Net. The forward pass is defined as $\mathbf{X}^l \to \mathbf{Z}^l \to \mathbf{P}^l \to \mathbf{X}^{l+1}$, where we calculate the NMSE with the updated parameters. The backward pass is defined as $\mathbf{X}^{l+1} \to \mathbf{P}^l \to \mathbf{Z}^l \to \mathbf{X}^l$, where we calculate the gradient of NMSE with respect to the parameters in each layer [20].

IV. Experiments

The performance of the proposed ADMM-Net framework is studied in the following experiments. The temperature data of a sensor network is collected by 54 sensors distributed in Intel Berkeley Research Lab [21] every 30 seconds between February 28, 2004 and April 5, 2004. The sensing data is regarded as the ground truth and is corrupted by Gaussian noise. The dimension of the measurement matrix $\mathbf{X}$ for network training and test is 49×60. The sampling rate is denoted as $r = \sum\mathbf{Q}(i,j)/MN$. We design a neural network with $L = 5$ in the following experiments.

The loss of ADMM-Net in network training with different learning rates is shown in Fig. 3, where the sampling rate

$r = 0.4$, noise power $\sigma^2 = 1$. The measurement data between 10:18 on February 29 and 12:18 on March 2 is selected for network training. As shown in Fig. 3, the loss decreases under different learning rates as the training step increases. The larger the learning rate is, the faster the NMSE converges, however, the greater the fluctuation is. The balance between convergence speed and accuracy can be achieved by selecting an appropriate learning rate.

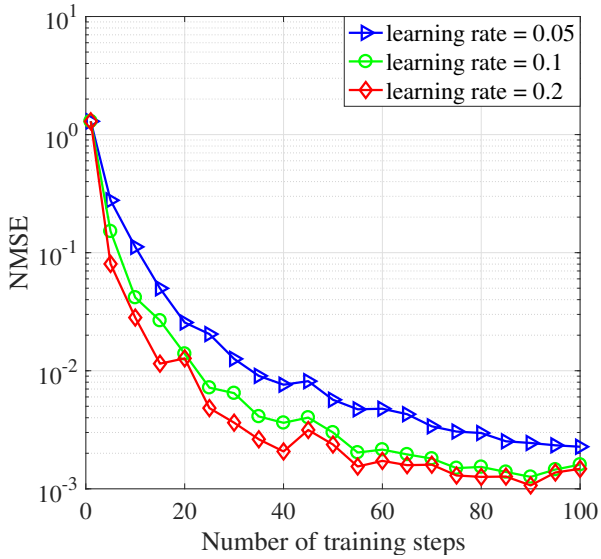

Fig. 3. Loss of ADMM-Net in network training.

Next, we would like to examine the recovery performance of ADMM-Net compared with other methods. The measurement data between 12:18 and 22:18 on March 2 is selected as the test dataset. The recovery accuracy of ADMM-Net, ADMM, LRDS, STCDG, and LRMA is shown in Fig. 4, where the sampling rate $r = 0.4$, noise power $\sigma^2 = 1$. In ADMM, $\lambda_1 = \lambda_2 = \rho = 0.001 \times 0.99^l$, $\tau = 0.8$, $\eta = 1$, $\mathbf{S}$ is built based on the location of nodes and complete sensing data. It can be observed that the ADMM-Net, LRDS, and ADMM achieve higher accuracy compared with other algorithms. The LRDS method performs basically the same as the ADMM, because both of them are solved by typical ADMM iterative algorithm. The ADMM-Net with learned parameters needs only $L = 5$ iterations while the ADMM takes about 62 iterations when the recovery accuracy is 2.5×10^{-3}. Thus, the proposed ADMM-Net can reduce a large number of iterations with slight accuracy loss.

The NMSE performance of the proposed ADMM-Net with different sampling rates and noise powers is shown in Fig. 5. From Fig. 5, we observe that the larger the sampling rate, the lower the NMSE, while larger sampling rate means more energy consumption of sensor nodes. Appropriate sampling rate should be chosen to achieve the trade-off between recovery accuracy and network lifetime. Noise power is another critical factor affecting data recovery performance. The proposed ADMM-Net achieves higher recovery accuracy when noise power is smaller, while the performance improvement is not significant when sampling rate $r \leq 0.1$, since sampling rate r becomes the main factor affecting performance in this case.

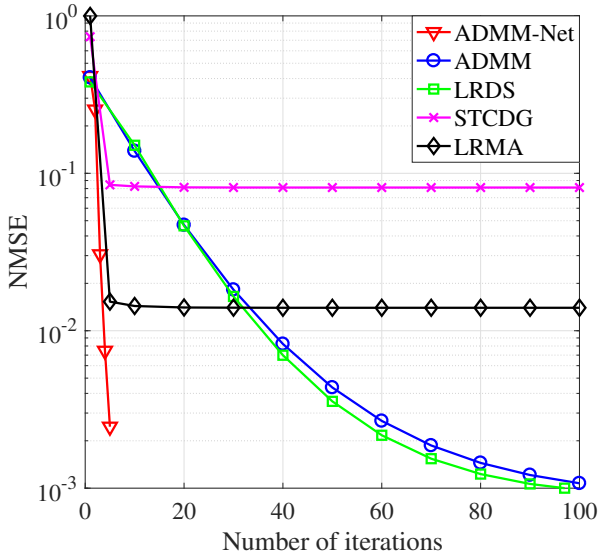

Fig. 4. Recovery performance of ADMM-Net.

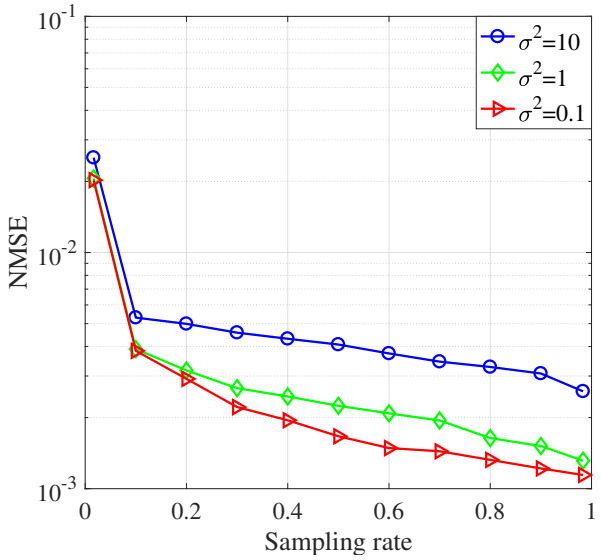

Fig. 5. Recovery performance of ADMM-Net with different sampling rates and noise powers.

V. Conclusions

In this paper, an ADMM-Net framework for correlated spatio-temporal data recovery was explored. We considered the spatial and temporal correlations of sensing data and adopted the ADMM algorithm to solve the convex optimization problem. Then we unrolled the typical ADMM algorithm into a fixed-length neural network that does not need to know the location of nodes. The experimental results validated that the proposed method can accelerate the convergence speed significantly with slight loss of recovery accuracy.

REFERENCES

[1] S. Ji, R. Beyah, and Z. Cai, "Snapshot and continuous data collection in probabilistic wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 3, pp. 626–637, Mar. 2014.

[2] X. Mao, K. Qiu, T. Li, and Y. Gu, "Spatio-temporal signal recovery based on low rank and differential smoothness," *IEEE Transactions on Signal Processing*, vol. 66, no. 23, pp. 6281–6296, Dec. 2018.

[3] M. El-Telbany and M. Maged, "Exploiting sparsity in wireless sensor networks for energy saving: A comparative study," *International Journal of Applied Engineering Research*, vol. 12, no. 4, pp. 452–460, 2017.

[4] J. Cheng, Q. Ye, H. Jiang, D. Wang, and C. Wang, "STCDG: An efficient data gathering algorithm based on matrix completion for wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 850–861, Feb. 2013.

[5] X. Piao, Y. Hu, Y. Sun, B. Yin, and J. Gao, "Correlated spatio-temporal data collection in wireless sensor networks based on low rank matrix approximation and optimized node sampling," *Sensors*, vol. 14, pp. 23137–23158, Dec. 2014.

[6] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of International Conference on Machine Learning (ICML)*, 2010, pp. 399–406.

[7] Y. Chen, Wei Yu, and T. Pock, "On learning optimized reaction diffusion processes for effective image restoration," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 5261–5269.

[8] O. Solomon, R. Cohen, Y. Zhang, Y. Yang, Q. He, J. Luo, R. J. van Sloun, and Y. C. Eldar, "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Transactions on Medical Imaging*, pp. 1–13, 2019, early access.

[9] P. Ochs, R. Ranftl, T. Brox, and T. Pock, "Techniques for gradient-based bilevel optimization with non-smooth lower level problems," *Journal of Mathematical Imaging and Vision*, vol. 56, no. 2, pp. 175–194, Oct. 2016.

[10] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *arXiv preprint arXiv:1912.10557*, 2019.

[11] R. Cohen, Y. Zhang, O. Solomon, D. Toberman, L. Taieb, R. J. van Sloun, and Y. C. Eldar, "Deep convolutional robust PCA with application to ultrasound imaging," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 3212–3216.

[12] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, "Efficient and interpretable deep blind image deblurring via algorithm unrolling," *IEEE Transactions on Computational Imaging*, pp. 1–16, 2020, early access.

[13] Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–18, Nov. 2018.

[14] X. Piao, Y. Hu, Y. Sun, and B. Yin, "Efficient data gathering in wireless sensor networks based on low rank approximation," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, Aug. 2013, pp. 699–706.

[15] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, Apr. 2009.

[16] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Mathematical Programming*, vol. 128, pp. 321–353, 2011.

[17] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Advances in Neural Information Processing Systems*, 2011, pp. 612–620.

[18] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[19] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Advances in Neural Information Processing Systems*, 2016, pp. 10–18.

[20] Y. Li, L. Huang, Y. Yin, Y. Wang, and G. Gui, "ADMM-Net for robust compressive sensing image reconstruction in the presence of symmetric α-stable noise," in *Proceedings of APSIPA Annual Summit and Conference*, 2018, pp. 296–300.

[21] Intel Berkeley Research Lab, "Intel Lab Data," http://db.lcs.mit.edu/labdata/labdata.html, 2004.