FAST DECENTRALIZED LEARNING VIA HYBRID CONSENSUS ADMM

Meng Ma[1,2]*, Athanasios N. Nikolakopoulos*[1]*, Georgios B. Giannakis*[1,2]

Digital Technology Center[1] and Dept. of Electrical and Computer Engineering[2], University of Minnesota
{maxxx971,anikolak,georgios}@umn.edu

ABSTRACT

The Alternating Directions Methods of Multipliers (ADMM) has witnessed a resurgence of interest over the past few years fueled by the ever increasing demand for scalable optimization techniques to tackle real-world statistical learning problems. However, despite its success in several application settings the applicability of the traditional centralized ADMM is limited by its communication requirement to a global *fusion center*, which might not be always feasible. Its decentralized variant D-CADMM, on the other hand, while it alleviates this need, it does so at the expense of significantly slower convergence in cases of adverse underlying network topologies. To address the aforementioned limitations, in this work we consider the presence of *multiple fusion centers* and we propose a unifying framework that allows leveraging the structure of the communication network to accelerate the decentralized ADMM even in cases where it is not practical to resort to its fully centralized counterpart. We prove the linear convergence rate of the proposed approach and we verify its promising performance by carrying out numerical tests on both real and synthetic networks.

Index Terms— Distributed Optimization, ADMM, Decentralized Learning, Consensus, Multi-agent Network

1. INTRODUCTION

Recent advances in Machine Learning and Data Mining have led to the formulation of increasingly demanding optimization problems. Thankfully, there exist certain commonly occurring cases that exhibit special structural characteristics (eg. separability, large sums of component functions, etc.) that make these problems amenable to parallel and distributed processing, thereby ensuring the applicability of the underlying methods in real world settings.

A simple and particularly well-suited approach for such optimization problems is the Alternating Direction Method of Multipliers (ADMM) [1, 2], which adopts a *multi-agent* optimization formulation where each agent iteratively updates its own parameters based on locally acquired data as well as information obtained from a global *Fusion Center* (FC). The convergence of ADMM has been studied extensively for long [1, 3], but it was not until recently that the linear convergence rate was established [4, 5].

To accommodate situations where completely decentralized learning is preferable, as well as to address cases where the communication to a global fusion center might not be feasible (e.g. due to privacy constrains or because of prohibitive communication cost, etc.), one can resort to *Decentralized Concensus ADMM*. Contrary to centralized ADMM, D-CADMM assumes that each agent can exploit only the information gathered from single-hop neighbors, and it has been successfully employed in several application [6, 7, 8, 9, 10].

1.1. Motivation & Contributions

In real world scenarios, the communication between the agents in D-CADMM is restricted to abide by the topology of an actual underlying network; a fact that can potentially impose severe limitations to the convergence rate of D-CADMM especially as the network's size increases. This is in part due to the inherent "single-hop philosophy" of D-CADMM, that does not allow it to exploit coarse-level structural properties of the underlying networks. However, large-scale networks arising in diverse application settings are known to exhibit nontrivial structural characteristics [11, 12, 13] as well as rich innate hierarchical organization [14, 15]. Motivated by this, here we try to address the following question: *Is it possible to exploit the structure of the communication network to accelerate D-CADMM?* Towards this goal, in this paper we propose a novel distributed learning framework, called **Hybrid Consensus ADMM** (H-CADMM). H-CADMM provides a unified view of ADMM-based distributed multi-agent optimization, naturally generalizing centralized ADMM and D-CADMM. Adopting an expressive modeling approach based on hypergraphs, H-CADMM allows for multiple *FC roles*, that can be assigned judiciously leveraging information about the communication network structure. Building on the results of Shi *et al.* [16] we establish linear convergence of H-CADMM and we showcase its potential based on numerical tests using both synthetic and real networks.

Notation. Vectors are represented by bold lower case, and matrices bold upper case. $\mathbf{I}_N$ denotes identity matrix of size $N \times N$. $\mathbf{1}(\mathbf{0})$ is a all one(zero) vector of appropriate size.

2. PRELIMINARIES

We model communication constraints as an undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, whose nodes correspond to *agents* and edges represent communication between nodes. We use N, M to denote the total number of nodes and edges, respectively, and $\mathcal{N}_i := \{j|(i,j) \in \mathcal{E}\}$ to denote neighbors of node i. Throughout this paper, we will make the following assumptions about the graph and the local objective functions.

Assumption 1 The graph is connected.

Assumption 2 Local objective function $f_i(\cdot)$ is differentiable and has Lipschitz continuous gradient, i.e. ,$\|\nabla f_i(x) - f_i(y)\|_2 \leq L\|x-y\|_2$.

Assumption 3 Local objective function $f_i(\cdot)$ is σ-strongly convex, i.e., for any x, y, $f_i(y) \geq f_i(x) + \nabla f_i(x)^\top (y-x) + \frac{\sigma}{2}\|x-y\|_2^2$.

2.1. Centralized consensus ADMM

In the centralized formulation, every node maintains a local estimate $\mathbf{x}_i$ of the global variable $\mathbf{x}$, and consensus is guaranteed by enforcing agreement of all agents on an global auxiliary variable. Thus, one can formulate the centralized constrained problem

$$\min_{\{x_i\}} \sum_{i=1}^{N} f_i(x_i), \quad \text{subject to } x_i = z. \tag{1}$$

Applying ADMM boils down to iteratively updating primal and dual variables as

$$x_i^{k+1} = (\nabla f_i + I)^{-1}(c\bar{x}^k - \lambda_i^k), \tag{2}$$

$$\lambda_i^{k+1} = \lambda_i^k + c(x_i^{k+1} - \bar{x}^k), \tag{3}$$

where $z = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ has been eliminated for simplicity, and we use (2) as a shorthand for x_i^{k+1} being the solution of $\nabla f_i(x_i^{k+1}) + x_i^{k+1} = (c\bar{x}^k - \lambda_i^k)$.

2.2. Decentralized consensus ADMM

In case no global *Fusion Center* (FC) is available, consensus can be achieved by enforcing agreement among neighbors, provided that the underlying graph is connected. Introducing auxiliary variables $\{z_{ij}\}_{(i,j)\in\mathcal{E}}$, we can reformulate (1) in decentralized form

$$\min_{\{x_i\}} \sum_{i=1}^{N} f_i(x_i), \quad \text{s.t. } x_i = z_{ij}, x_j = z_{ij}, (i,j) \in \mathcal{E}. \tag{4}$$

Following the steps described in [16], ADMM updates entail

$$x_i^{k+1} = (\nabla f_i + c|\mathcal{N}_i|I)^{-1}\left(\frac{c}{2}\sum_{j\in\mathcal{N}_i}(x_i^k + x_j^k) - \alpha_i^k\right) \tag{5}$$

$$\alpha_i^{k+1} = \alpha_i^k + \frac{c}{2}\sum_{j\in\mathcal{N}_i}(x_i^{k+1} - x_j^{k+1}). \tag{6}$$

3. HIERARCHICAL CONSENSUS ADMM

In this paper, we are considering the case when there are no global FCs but multiple local FCs in the network. This situation is pretty common in large-scale networks since it's not feasible to deploy only one global FC, considering, for instance, heavy communication burden of the FC. Neither of the aforementioned methods can deal with this situation as none of them is capable of handling *hybrid constraints*, i.e. the need of some nodes to exchange information with both local FCs and its single-hop neighbors at the same time. It turns out *hybrid constraints* can be precisely captured by *hypergraphs*, in which a local FC is modeled as one hyperedge containing all its connected nodes, while a simple edge connecting only two nodes is modeled as one hyperedge containing two nodes.

3.1. Problem formulation

Assume an ordering of the hyperedges $1, 2, \ldots, M$ and introduce auxiliary variable z_j for hyperedge j, then *hybrid constraints* can be readily reparameterized as $x_i = z_j$. Suppose $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{z} \in \mathbb{R}^M$ are vectors collecting all local and edge variables, and assume the total number of constraints is T, then we can define $\mathbf{A} \in \mathbb{R}^{T\times N}$ and $\mathbf{B} \in \mathbb{R}^{T\times M}$ such that when the t-th constraint is $x_i - z_j = 0$, $A_{ti} = 1$ and $B_{tj} = 1$; all other elements of t-th rows are zero. With the help of newly defined variables, we formulate the following *hybrid constrained* problem in a compact form

$$\min_{x_i} \sum_{i=1}^{N} f_i(x_i), \quad \text{subject to } \mathbf{Ax} - \mathbf{Bz} = \mathbf{0}. \tag{7}$$

Furthermore, let $\mathbf{C} \in \mathbb{R}^{N\times M}$ be the incidence matrix of the hypergraph, that is, $C_{ij} = 1$ if node i is connected to edge j. Let the diagonal matrix $\mathbf{D} \in \mathbb{R}^{N\times N}$ be the node degree matrix, i.e., D_{ii} stores the degree of node i; let diagonal matrix $\mathbf{E} \in \mathbb{R}^{M\times M}$ represent "edge degree" matrix, whose diagonal element E_{jj} is the degree of hyperedge j, i.e., number of nodes connected to hyperedge j. If all the constraints are ordered increasingly according to node labels and edge labels, one can verify

$$\mathbf{A}^\top\mathbf{A} = \mathbf{D}, \quad \mathbf{B}^\top\mathbf{B} = \mathbf{E}, \quad \mathbf{A}^\top\mathbf{B} = \mathbf{C}. \tag{8}$$

3.2. Algorithm

Applying ADMM to (7) involves solving following equations

$$\nabla f(\mathbf{x}^{k+1}) + \mathbf{A}^\top\boldsymbol{\lambda}^k + c\mathbf{A}^\top(\mathbf{Ax}^{k+1} - \mathbf{Bz}^k) = 0, \tag{9}$$

$$\mathbf{B}^\top\boldsymbol{\lambda}^k + c\mathbf{B}^\top(\mathbf{Ax}^{k+1} - \mathbf{Bz}^{k+1}) = 0, \tag{10}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + c(\mathbf{Ax}^{k+1} - \mathbf{Bz}^{k+1}). \tag{11}$$

Left multiplying (11) by $\mathbf{B}^\top$ and add it to (10) yields

$$\mathbf{B}^\top\boldsymbol{\lambda}^{k+1} = \mathbf{0}. \tag{12}$$

Algorithm 1: Hybrid consensus ADMM

1: Initialize $\mathbf{x}^0$, $\boldsymbol{\alpha}^0$, ϵ
2: **while** primal or dual residual greater than ϵ **do**
3: update $\mathbf{x}$ at each node according to (15)
4: update $\mathbf{z}$ at each local FC according to (13)
5: update $\boldsymbol{\alpha}$ at each node according to (14)
6: **end while**
7: **return** $\mathbf{x}$

Assume $\boldsymbol{\lambda}$ is initialized such that $\mathbf{B}^\top\boldsymbol{\lambda}^0 = 0$, then it can be guaranteed (12) will always hold for all $k \geq 0$. Consequently, we can solve (10) for $\mathbf{z}$

$$\mathbf{z}^{k+1} = \mathbf{E}^{-1}\mathbf{C}^\top\mathbf{x}^{k+1}. \tag{13}$$

Similarly, left multiply (11) by $\mathbf{A}^\top$ and introduce a new variable $\boldsymbol{\alpha}^k = \mathbf{A}^\top\boldsymbol{\lambda}^k$

$$\boldsymbol{\alpha}^{k+1} - \boldsymbol{\alpha}^k = c(\mathbf{D}\mathbf{x}^{k+1} - \mathbf{C}\mathbf{z}^{k+1}). \tag{14}$$

Plugging (13) into (9) yields

$$\mathbf{x}^{k+1} = (\nabla f + c\mathbf{D}I)^{-1}(c\mathbf{C}\mathbf{z}^k - \boldsymbol{\alpha}^k). \tag{15}$$

We are assuming a scalar version of this algorithm to simplify the notation as well as analysis. It's trivial to generalize to vector case by extending corresponding matrices to block structure.

3.3. Connections to C-CADMM and D-CADMM

Modeling constraints as a hypergraph affords great flexibility, and provides a unified view of consensus ADMM formulations. Indeed it is not difficult to see both centralized and decentralized versions are special cases of our proposed approach.

C-CADMM as a special case: There is only one global FC, therefore $\mathbf{A} = \mathbf{I}$, $\mathbf{B} = \mathbf{1}, M = 1$, consequently, $\mathbf{D} = \mathbf{A}^\top\mathbf{A} = \mathbf{I}$, $\mathbf{E} = \mathbf{B}^\top\mathbf{B} = N$, $\mathbf{C} = \mathbf{A}^\top\mathbf{B} = \mathbf{1}$. Then (13) is actually $z^{k+1} = \frac{1}{N}\sum_{i=1}^N x_i^k = \bar{x}^k$. As a result, (15) and (14) are in fact

$$\mathbf{x}^{k+1} = (\nabla f + cI)^{-1}(c\mathbf{1}\bar{x}^k - \boldsymbol{\lambda}^k), \tag{16}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + c(\mathbf{x}^{k+1} - \mathbf{1}\bar{x}^{k+1}), \tag{17}$$

where we have used the fact $\boldsymbol{\alpha}^k = \mathbf{A}^\top\boldsymbol{\lambda}^k = \boldsymbol{\lambda}^k$. Comparing (16) to (2), one realize (2) is indeed the entry-wise update. Similar observation holds for (17) and (3).

D-CADMM as a special case: In this case, each hyperedge has a degree of 2. Therefore, $\mathbf{E} = 2\mathbf{I}$, and (13) now becomes

$$\mathbf{z}^{k+1} = \frac{1}{2}\mathbf{C}^\top\mathbf{x}^{k+1}. \tag{18}$$

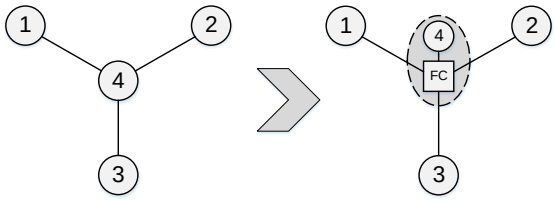

Fig. 1. Illustration of accelerating D-CADMM

Using (18) to eliminate $\mathbf{z}$, (15) and (14) can be simplifed as

$$\mathbf{x}^{k+1} = (\nabla f + c\mathbf{D}I)^{-1}(\frac{c}{2}\mathbf{C}\mathbf{C}^\top\mathbf{x}^k - \boldsymbol{\alpha}^k), \tag{19}$$

$$\boldsymbol{\alpha}^{k+1} = \boldsymbol{\alpha}^k + c(\mathbf{D} - \frac{1}{2}\mathbf{C}\mathbf{C}^\top)\mathbf{x}^{k+1}. \tag{20}$$

Noticing $D_{ii} = |\mathcal{N}_i|$ and $(\mathbf{C}\mathbf{C}^\top\mathbf{x})_i = \sum_{j\in\mathcal{N}_i}(x_i + x_j)$, one realizes (5) and (6) are the entry-wise version of (19) and (20).

4. LINEAR CONVERGENCE RATE

Consider Algorithm 1 for solving hybrid consensus optimization problems and introduce

$$\mathbf{u} = \begin{bmatrix}\mathbf{z}\\ \boldsymbol{\lambda}\end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix}c\mathbf{E} & \mathbf{0}\\ \mathbf{0} & \frac{1}{c}\mathbf{I}_T\end{bmatrix},$$

where $\mathbf{u}$ is the concatenation of primal variable $\mathbf{z}$ and dual variable $\boldsymbol{\lambda}$. Obviously, $\mathbf{G}$ is positive definite, based on which we can define "$\mathbf{G}$-norm" $\|\mathbf{x}\|_\mathbf{G} = \mathbf{x}^\top\mathbf{G}$, for $\mathbf{x} \in \mathbb{R}^{M+T}$. Thanks to the strong convexity and Lipschitz continuous gradient, we can show that $\mathbf{u}^k$ is Q-linearly convergent to its optimal value $\mathbf{u}^\star$ in the $\mathbf{G}$-norm sense. Furthermore, the Q-linear convergence of $\mathbf{u}^k$ implies the R-linear convergence of $\mathbf{x}^k$ to its optimal value $\mathbf{x}^\star$.

Theorem 1. *Consider the ADMM updates in* (13)(14)(15)*. If the objective function is σ-strongly convex and its gradient is L-Lipschitz continuous, then for any $\mu > 1$, $\mathbf{u}^k$ is Q-linearly convergent to its optiml value $\mathbf{u}^\star$ with respect to G-norm*

$$\|\mathbf{u}^{k+1} - \mathbf{u}^\star\|_\mathbf{G}^2 \leq \frac{1}{1+\delta}\|\mathbf{u}^k - \mathbf{u}^\star\|_\mathbf{G}^2,$$

where the constant δ satisfies the inequality

$$\delta \leq \min\left\{\frac{\mu-1}{\mu}\frac{\lambda_{min}(\mathbf{D})}{\lambda_{max}(\mathbf{D})}, \frac{2\sigma}{c\lambda_{max}(\mathbf{C}\mathbf{E}^{-1}\mathbf{C}^\top) + \frac{\mu L^2}{c\lambda_{min}(\mathbf{D})}}\right\}.$$

Furthermore, $\mathbf{x}^k$ is R-linearly convergent to $\mathbf{x}^\star$ in the sense that

$$\|\mathbf{x}^{k+1} - \mathbf{x}^\star\|_2^2 \leq \frac{1}{m}\|\mathbf{u}^k - \mathbf{u}^\star\|_\mathbf{G}^2.$$

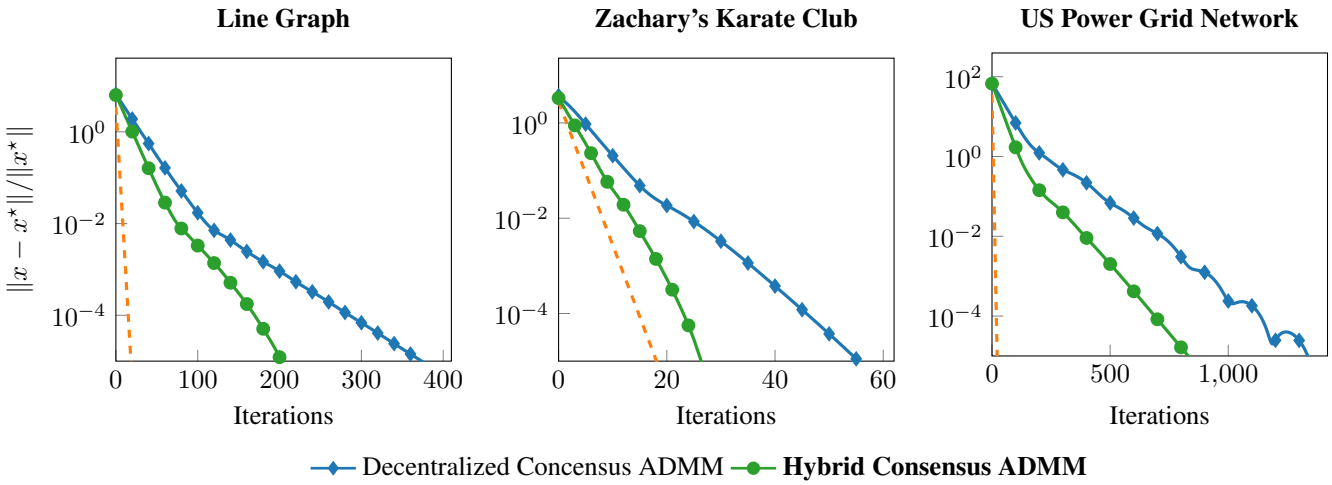

Fig. 2. Relative accuracy vs number of iterations of H-CADMM and D-CADMM on: (a) a line graph with 50 nodes and 49 edges (left plot); (b) Karate club network with 34 nodes 76 edges (middle plot); (c) US power grid network with 4,941 nodes and 6,594 edges (right plot). For reference we also include the performance of the centralized version (dashed line).

5. FAST DECENTRALIZED LEARNING

Decentralized learning is ideal for parallel processing when global FC is not available for various concerns, since it only involves information exchange among neighbors. However, it also suffers from slow convergence – especially when the diameter of the underlying graph is large; a fact that may impede its applicability to certain real-world settings.

Thankfully, the framework we develop in this work can alleviate this problem. In particular, our analysis suggests that one can accelerate the convergence of D-CADMM simply by denoting a carefully selected subset of nodes on the underlying graph as local *virtual-FCs* and proceed by applying our proposed algorithm. A prominent advantage of this approach is that there is no need to have a dedicated FC, nor to modify existing edges. The creation of each virtual FC can be done simply by selecting a node and then connecting the node itself and all its edges to the newly created virtual FC. The procedure is illustrated in Fig. 1.

The issue of selecting the the right nodes for the job is non trivial and it is intuitively apparent that it should be governed by the properties of the underlying graph. Here, for simplicity, we use the nodes with the highest degree.

6. NUMERICAL TESTS

In this section, we perform several numerical experiments to test our algorithm using both synthetic and real networks. For all the experiments, we assume that the agents try to measure x_0 based on observation $y_i = x_0 + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, 0.1)$. Each agent tries to minimize estimation error $f_i(x) = \frac{1}{2}\|y_i - x\|_2^2$. We solve the same problem on different networks using C-CADMM, D-CADMM with standard setting and H-CADMM using the method depicted in Sec. 5 by setting nodes with degree greater than 2 as local *virtual-FCs*. Penalty parameter c is tuned individually for best performance. The relative accuracy, $\|x^k - x^\star\|_2/\|x^\star\|_2$ vs. iteration number is recorded for comparison.

We first test on a line graph with 50 nodes. Line graphs have the largest diameter given the number of nodes, thus, they represent the worst case for D-CADMM. Simulation results, presented in Fig. 2, show that H-CADMM performs markedly better than D-CADMM, demonstrating our approach can achieve considerable acceleration even for the most difficult cases.

The second experiment is performed on Zachary's karate club [17]. This network captures 34 members of a karate club as nodes, and their interaction outside the club as edges. Visually one can discover many clusters in this network, suggesting potential acceleration. Indeed, results in Fig. 2 show a large gap between H-CADMM and D-CADMM.

Finally, we test our algorithm on a large real network, the US power grid [18]. This network consists of 4,941 nodes, representing generators, transformers or substations, 6,594 edges, representing high-voltage transmission line. The results are presented in Fig. 2. Though it takes much more iterations to achieve the same accuracy, which is not surprising given the size of network, our approach does accelerate D-CADMM by a considerable margin.

7. CONCLUSION

In this paper, we proposed the novel H-CADMM that generalizes both centralized and decentralized consensus ADMM in a unified view, enabling the acceleration of D-CADMM with modified udpates. We also provide linear convergence rate H-CADMM. Numerical tests demonstrate its effectiveness.

8. REFERENCES

[1] Dimitri P. Bertsekas and John N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, vol. 23, Prentice hall Englewood Cliffs, NJ, 1989.

[2] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan 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.

[3] Tom Goldstein, Brendan O'Donoghue, Simon Setzer, and Richard Baraniuk, "Fast alternating direction optimization methods," *SIAM Journal on Imaging Sciences*, vol. 7, no. 3, pp. 1588–1623, 2014.

[4] Mingyi Hong and Zhi-Quan Luo, "On the Linear Convergence of the Alternating Direction Method of Multipliers," *arXiv preprint*, Aug. 2012.

[5] Wei Deng and Wotao Yin, "On the Global and Linear Convergence of the Generalized Alternating Direction Method of Multipliers," *Journal of Scientific Computing*, vol. 66, no. 3, pp. 889–916, May 2015.

[6] Pedro A. Forero, Alfonso Cano, and Georgios B. Giannakis, "Distributed feature-based modulation classification using wireless sensor networks," in *MILCOM 2008-2008 IEEE Military Communications Conference*. 2008, pp. 1–7, IEEE.

[7] Pedro A. Forero, Alfonso Cano, and Georgios B. Giannakis, "Convergence analysis of consensus-based distributed clustering.," in *ICASSP*. 2010, pp. 1890–1893, Citeseer.

[8] G. Mateos, J. A. Bazerque, and G. B. Giannakis, "Distributed Sparse Linear Regression," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5262–5276, Oct. 2010.

[9] Vassilis Kekatos and Georgios B. Giannakis, "Distributed robust power system state estimation," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, 2013.

[10] Georgios B. Giannakis, Qing Ling, Gonzalo Mateos, Ioannis D. Schizas, and Hao Zhu, "Decentralized Learning for Wireless Communications and Networking," in *Splitting Methods in Communication, Imaging, Science, and Engineering*, Scientific Computation, pp. 461–497. Springer, Cham, 2016.

[11] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos, "On power-law relationships of the internet topology," *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 4, pp. 251–262, Aug. 1999.

[12] Athanasios N Nikolakopoulos and John D Garofalakis, "Ncdawarerank: a novel ranking method that exploits the decomposable structure of the web," in *Proceedings of the sixth ACM international conference on Web search and data mining*. ACM, 2013, pp. 143–152.

[13] Mark EJ Newman, "The structure and function of complex networks," *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

[14] Athanasios N. Nikolakopoulos and John D. Garofalakis, "On the multi-level near complete decomposability of a class of multiprocessing systems," in *Proc. of the 20th PCI*, New York, NY, USA, 2016, PCI '16, pp. 14:1–14:6, ACM.

[15] Santo Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3, pp. 75–174, 2010.

[16] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, "Linearly convergent decentralized consensus optimization with the alternating direction method of multipliers," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 4613–4617.

[17] Wayne W. Zachary, "An Information Flow Model for Conflict and Fission in Small Groups," *Journal of Anthropological Research*, vol. 33, no. 4, pp. 452–473, 1977.

[18] Duncan J. Watts and Steven H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, June 1998.