ADMM and Random Walks on Graphs

Guilherme França
Boston College, Johns Hopkins University
guifranca@gmail.com

José Bento
Boston College
jose.bento@bc.edu

Abstract

A connection between distributed ADMM and lifted Markov chains was recently proposed for a non-strictly-convex consensus problem parametrized by a graph [1]. This was followed by a conjecture that ADMM is faster than GD by a square root factor in its convergence time, in close analogy to the mixing speedup achieved by lifting several Markov chains. Nevertheless, a proof is still lacking. Here we fully characterize distributed over-relaxed ADMM for this same problem in terms of the topology of the graph. More specifically, we relate its convergence rate with the mixing time of random walks on the graph. A consequence of our results is a proof of the aforementioned conjecture, which, interestingly, is valid for any graph, even those whose random walks cannot be accelerated via Markov chain lifting.

1 Introduction

Many problems in statistics and machine learning involve solving an optimization problem. In this current age of ever increasing datasets, traditional in-memory methods do not scale, so distributed algorithms play a fundamental role. The Alternating Direction Method of Multipliers (ADMM) is one such excellent example, since it is extremely versatile, e.g. does not require differentiability of the objective, it is often easy to implement and it is easily distributed [2].

In this paper we analyze how the *exact and optimally tuned* asymptotic convergence rate of a distributed implementation of over-relaxed ADMM depends on the topology of an underlying network, through which several agents solve local problems and share messages with the goal of solving an optimization problem. We define the asymptotic convergence rate, τ, of an algorithm by

$$\log \tau \equiv \lim_{t\to\infty} \max_{\|\boldsymbol{z}^0\|\leq 1} \left\{ t^{-1} \log \|\boldsymbol{z}^\star - \boldsymbol{z}^t\| \right\}, \tag{1}$$

where $t = 0, 1, \ldots$ is the iteration time and $\boldsymbol{z}^t$ is the current estimate of the solution, which is assumed to converge to $\boldsymbol{z}^\star$. The optimally tuned asymptotic rate of convergence, τ^*, is the minimum of τ over algorithm's parameters.

Our over-relaxed ADMM scheme is the same as in [1] and it depends on the over-relaxed parameter $0 < \gamma < 2$ and the penalty parameter $\rho > 0$. We also compare ADMM with gradient descent (GD) which depends on the step-size parameter $\alpha > 0$. Note that GD can also be implemented in a distributed way: each agent computes the gradient of part of the objective, sharing it with other agents that need this gradient to update their local copies of z_t.

Given an optimization problem parameterized by a network $\mathcal{G}$, one of our motivations is to understand whether ADMM is more or less sensitive to the topology of $\mathcal{G}$ than GD. In other words, if τ_A and τ_G are the convergence rates of ADMM and GD, respectively, we want to compare $\tau_A^\star(\mathcal{G})$ and $\tau_G^\star(\mathcal{G})$. Another motivation comes from an interesting idea first presented at NIPS OPT2016 Workshop [1] which relates distributed ADMM to *lifted* Markov chains [3]. Based on a lifting relation it was conjectured that

$$1 - \tau_A^\star \geq C\sqrt{1 - \tau_G^\star} \tag{2}$$

for some constant C. The inequality (2) was supported by empirical evidence but its proof was lacking, since for the problem considered in [1] existing bounds on the convergence rate could not be

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

directly applied. Here we aim to solve this conjecture and, in particular, show that (2) is valid for *any* graph $\mathcal{G}$, even graphs whose Markov chains do not accelerate via lifting.

Consider an undirected, connected, and simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ the set of edges. We study the following consensus problem over $\mathcal{G}$ (was also studied in [1]):

$$\min_{\boldsymbol{z} \in \mathbb{R}^{|\mathcal{V}|}} \left\{ f(\boldsymbol{z}) = \frac{1}{2} \sum_{(i,j) \in \mathcal{E}} (z_i - z_j)^2 \right\}. \tag{3}$$

The reason behind (3) is to make the analysis independent of properties of the objective function other than $\mathcal{G}$, such as its curvature or scaling factors. Note that the communication network over which distributed ADMM and GD solve (3) is essentially equal to $\mathcal{G}$ itself. Problem (3) is deceptively simple; it has trivial solution $x_i = x_j$ if i and j belong to the same connected component of $\mathcal{G}$. However, it is not immediately obvious to which of these infinitely many possible solutions a given distributed algorithm will converge to. Different agents of a distributed algorithm have to communicate to agree on a final solution, and the speed at which they reach consensus is a non-trivial problem. For instance, if we solve (3) through ADMM we have one agent per term of the objective function, and each agent has local copies of all the variables involved. The final solution is a vector where each component equals the average of the initial values of these local variables. Therefore, unsuspectingly, we have solved a non-trivial distributed-average consensus problem, although in different form than typically studied; see e.g. [4, 5]. Moreover, the objective function (3) naturally appears in several interesting problems. A classical example is the graph interpolation problem [6], where one solves (3) subject to $z_i = c_i$ for $i \in \mathcal{V}'$ where $\mathcal{V}' \subset \mathcal{V}$ and c_i is a fixed constant. Our analysis of ADMM applied to (3) may provide insights into graph interpolation and other important problems.

2 Related Work

Our results cannot be derived from existing literature on the convergence of ADMM. First, most results only compute non-optimal upper bounds on the global convergence, as opposed to (1). Second, our convergence rate for (3) is linear, while most results with linear rates assume strong convexity; see e.g. [7, 8, 9, 10]. Linear rates are proven without strong convexity in [11], but they are not tight enough for our purposes. Many results avoid strong convexity by focusing on the convergence rate of the objective rather than the iterates; see e.g. [12, 13].

The authors of [4, 5] study ADMM for the consensus problem $f(\boldsymbol{z}) = \sum_{i \in \mathcal{V}} \sum (z_i - c_i)^2$, subject to $z_i = z_j$ if $(i, j) \in \mathcal{E}$, where $c_i > 0$ are constants. This problem is strongly convex and it is not equivalent to (3). Several papers consider $f(\boldsymbol{z}) = \sum_i f_i(\boldsymbol{z})$ with ADMM updates that are insensitive to whether or not $f_i(\boldsymbol{z})$ depends on a subset of the components of $\boldsymbol{z}$; see [14, 15, 16, 17, 18] and references therein. These papers solve $\min_{\mathbf{z}} \sum_i f_i(\boldsymbol{z})$ over a communication network by recasting this problem as minimizing $\sum_i f_i(\boldsymbol{x}_i)$ subject to $\boldsymbol{x}_i = \boldsymbol{x}_j$ if (i, j) are edges in the network. In our setting, distributed ADMM is a message-passing algorithm where the messages between agents i and j are related only to the variables shared by functions f_i and f_j. Thus, our ADMM scheme is different. Several of these works try to understand how topology affects convergence, for example [15, 16, 17, 18]. Among these, the results of [16, 17] are applicable to non-strongly-convex objectives but do they do not prove linear convergence rates for the variables we intend to optimize.

For quadratic problems, known ADMM rates and optimal parameters are more explicit/tight; see e.g. [19, 20, 21, 22, 23, 24, 25]. However, the required assumptions do not hold for the non-strongly convex problem (3). Moreover, there are very few results directly comparing τ_A^* and τ_G^*. An explicit comparison is provided in [26], but assumes strong convexity and their scheme is not fully distributed. The work [27] considers non-strongly-convex quadratic problems, and to the best of our knowledge this is the only work, such as we do here, analyzes ADMM in terms of complex eigenvalues of the recursion matrix. However, they do not provide optimal bounds in terms of the graph $\mathcal{G}$. The work [28] also study ADMM for non-strictly-convex quadratic problems, but their error rate is defined differently than (1) and it is not clear if both are comparable. Moreover, they provide generic bounds which are not optimally tuned to the case of problem (3).

Our work makes connections between ADMM, GD, and (lifted) Markov chains. Lifted Markov chains were previously employed to speedup convergence time of distributed averaging and gossip algorithms [29, 30, 31], however, these algorithms are not related to ADMM.

3 Main results

Due to lack of space we state our theorems in the case where $\mathcal{G}$ has at least one even cycle and conductance $\Phi \leq 1/2$, i.e. random walks on $\mathcal{G}$ mix slowly, which is the most interesting case. We have similar results for any choice of $\mathcal{G}$.

Theorem 1 (Optimal convergence rate of ADMM)**.** *Assume that the graph $\mathcal{G}$ has at least one cycle of even length, and conductance $\Phi \leq 1/2$. Let $\mathcal{W}$ be the transition matrix of a random walk on $\mathcal{G}$, and denote its second largest eigenvalue by $\omega^\star = \lambda_2(\mathcal{W})$. The best possible convergence rate of ADMM and its optimal parameters are given by*

$$\tau_A^\star = \gamma^\star - 1, \qquad \gamma^\star = \frac{4}{3 - \sqrt{(2-\rho^\star)/(2+\rho^\star)}}, \qquad \rho^\star = 2\sqrt{1-(\omega^\star)^2}. \tag{4}$$

One can show that (4) is always an upper bound for τ_A^* regardless of the topology of $\mathcal{G}$. Note that ω^* is directly related to the mixing time of a random walk on $\mathcal{G}$ [32]. We now give a sketch of the proof of Theorem 1. Full details will be presented in a longer version of this paper.

Proof sketch. From [1, Thm. 1] we know that the convergence rate of ADMM, τ_A, is equal to the second largest eigenvalue, in absolute value, of $T_A = I - \gamma(A + B - 2BA)$, where $A = (I + \rho^{-1}Q)^{-1}$ with $Q = \operatorname{diag}(\dots, Q_e, \dots)$ block diagonal, $Q_e = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for $e \in \mathcal{E}$, and $B = S(S^\top S)S^\top$ where S is a $2|\mathcal{E}| \times |\mathcal{V}|$ matrix such that, ordering the edges in $\mathcal{E}$, if the rth term in (3) depends on z_i, then $S_{2r-1,y} = S_{2r,y} = 1$ and zero otherwise. Note that the largest eigenvalue of T_A is one and is responsible for the non-strong-convexity and the infinite number of solutions of (3).

By direct substitution one can check that $T_A = (1 - \frac{\gamma}{2})I + \frac{\gamma}{\rho+2}U$, where $U = \Omega + \frac{\rho}{2}\widetilde{B}$, $\widetilde{B} = \widetilde{B}^\top = 2B - I$, $\Omega = \widetilde{B}R$, and $R = R^\top = I - Q$. In particular, Ω is orthogonal, $\widetilde{B}^2 = I$, and $R^2 = I$. The second largest eigenvalue (in absolute value) of T_A, and thus τ_A, can be determined once we know the spectrum of U. After we know τ_A, we can optimize it over γ and ρ to finish the proof.

Note that if $\rho = 0$ then $U = \Omega$ is orthogonal and its eigenvalues lie on the unit circle in the complex plane. We thus expect that, for ρ sufficiently small, the eigenvalues of U lie in a perturbation of this circle. It turns out that we can say much more than this.

Check by substitution that $U^{-1} = \left(1 - \rho^2/4\right)^{-1}\left(\Omega^\top - \frac{\rho}{2}\widetilde{B}\right)$. Defining $\Omega_S = \frac{1}{2}\left(\Omega + \Omega^\top\right)$ we have that $\Omega_S = \frac{1}{2}\left(U + (1 - \rho^2/4)U^{-1}\right)$. This last equation allows us to relate the spectrum of Ω_S with the spectrum of U. In particular, one can prove that, for all eigenvalues $\lambda(\Omega_S)$ not in $\{-1, 1\}$, $\lambda(\Omega_S) \pm i\sqrt{1 - \rho^2/4 - \lambda^2(\Omega_S)}$ is an eigenvalue of U. One can also prove that any eigenvalue of U must be of this form for some $\lambda(\Omega_S)$. Furthermore, it turns out that the spectrum of Ω_S is (with the exception of a few eigenvalues and different multiplicities) the same as the spectrum of the transition matrix $\mathcal{W}$. Therefore, we conclude that the eigenvalues $\lambda(T_A)$ and $\lambda(\mathcal{W})$ are, in general, related by

$$\lambda(T_A) = \left(1 - \frac{\gamma}{2}\right) + \frac{\gamma}{2+\rho}\left(\lambda(\mathcal{W}) \pm i\sqrt{1 - \frac{\rho^2}{4} - \lambda^2(\mathcal{W})}\right). \tag{5}$$

From this expression (and after computing a few other special eigenvalues) we find τ_A, hence we optimize it over γ and ρ. See Figure 1 (left) for an illustration. □

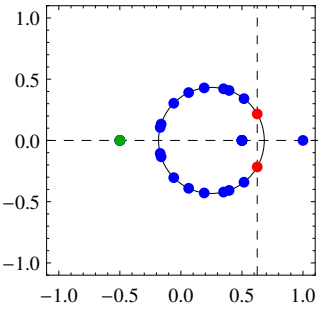
Our next theorem basically proves the conjecture of [1] relating $\tau_A^\star$ and $\tau_G^\star$.

Theorem 2 (ADMM speedup)**.** *Assume that $\mathcal{G}$ has an even length cycle and $\Phi \leq 1/2$, such that Theorem 1 holds. Then, there is $C = 1 - \mathcal{O}\left(\sqrt{\delta}\right)$ such that*

$$C\left(1 - \tau_G^\star\right) \leq \left(1 - \tau_A^\star\right)^2 \leq 2\Delta C\left(1 - \tau_G^\star\right), \tag{6}$$

where $\Delta = d_{\max}/d_{\min}$ is the ratio of the maximum to the minimum degree of $\mathcal{G}$. Here $\delta = 1 - \omega^\star$ is the spectral gap.

Notice that the upper bound in (6) implies that ADMM cannot improve much more than this square root factor. However, this upper bound becomes more loose for very irregular graphs, which have $\Delta \gg 1$, compared to regular graphs, which have $\Delta = 1$. Moreover, as mentioned before, since Theorem 1 provides an upper bound on $\tau_A^\star$ regardless of the topology of $\mathcal{G}$, the lower bound in (6) still remains valid for *any* graph.



$\mathcal{G}$	(a)	(b)	(c)	(d)
Φ	$1/3$	$1/2$	1	$1/5$
$\omega^\star$	$1/2$	$1/3$	$-1/3$	$\frac{\sqrt{97}-1}{12}$
$\rho^\star$	1.1732	1.886	2	1.351
$\gamma^\star$	1.464	1.414	$4/3$	1.659
$\tau_A^\star$	0.464	0.414	$1/3$	0.536
Thm. 1	✓	✓	✗	✗

Figure 1: *Left:* Illustration supporting Theorem 1. *Right:* Application of Theorem 1.

Proof sketch. If is straightforward to check that the convergence rate of GD is controlled by the second largest eigenvalue , in absolute value, of $I - \alpha L$, where L is the laplacian of $\mathcal{G}$. This implies that $\tau_G^* = \min_\alpha \max\left\{|1 - \alpha\,\bar{\ell}|, |1 - \alpha\,\ell^\star|\right\} = (\bar{\ell} - \ell^\star)/(\bar{\ell} + \ell^\star)$, where $\bar{\ell}$ is the largest eigenvalue of L and $\ell^\star$ is the second smallest and nonzero eigenvalue of L.

If τ_G^* was a function of ω^* we could eliminate ω^* and write τ_A^* as a function of τ_G^*. The situation is a little more complicated but we can get around this by relating ω^*, ℓ and $\bar{\ell}$ using of the following bounds [33, Lemmas 2.12 and 2.21]: $\ell^\star/d_{\max} \leq 1 - \omega^* \leq \ell^\star/d_{\min}$ and $d_{\max} \leq \bar{\ell} \leq 2d_{\max}$. The proof now follows from these bounds and the expressions for τ_G^* above and τ_A^* in (4). □

4 Numerical results

Figure 1 (left) supports the proof of Theorem 1. Over-relaxed ADMM for problem (3) evolves according to a linear system controlled by the eigenvalues of the transition matrix T_A. The proof boils down to relating the eigenvalues of T_A with the eigenvalues of the transition matrix $\mathcal{W}$ of the graph $\mathcal{G}$; see (5). In the Figure we plot the eigenvalues of T_A in the complex plane for a randomly generated graph $\mathcal{G}$. From (5) we see that for $\rho = 0$ the eigenvalues lie on a circle, with center $(1 - \gamma/2)$. As ρ increases each complex conjugate pair moves vertically towards the real-axis, shrinking the radius of the circle. When they reach it, they split left and right. To compute τ_A^*, we need to keep track of the eigenvalues marked in red and green, which are the second largest in absolute value (the largest is $\lambda(T_A) = 1$). The red eigenvalues are obtained from (5) with $\lambda(\mathcal{W})$ being the second largest eigenvalue of $\mathcal{W}$, and the green eigenvalue is equal to $1 - \gamma$. Finally, we need to choose ρ^* and γ^* such that the absolute value of the red eigenvalues equals the absolute value of the green eigenvalue when both are as small as possible. More specifically, we need to choose ρ^* such that the red eigenvalues have just touched the real-axis and choose γ^* such that $1 - \gamma^*$ equals the magnitude of the red eigenvalues.

We refer the reader to Figures 3, 4 and 5 from [1] for an illustration of Theorem 2. These Figures were used by [1] to support their conjecture, which we have now proved.

In the Table of Figure 1 (right) we illustrate two settings where Theorem 1 is valid, items (a) and (b), and two settings where we need analogs of Theorem 1 (which we will present in a longer version of this paper), items (c) and (d). The convergence rate τ_A^* was computed numerically and compared with Theorem 1.

5 Final Remarks

We have analyzed distributed over-relaxed ADMM for a consensus problem that cannot be directly solved using existing results. We were able to related the convergence rate of ADMM with random walks over its communication network. Our analysis also solved the conjecture proposed by [1] relating ADMM, GD and Markov chain lifting. In future work, we plan to search for other consensus problems, parametrized by a graph, for which analogous results hold.

Acknowledgements. This work is supported by NIH/1U01AI124302 and NSF/7433/8083 grants.

References

[1] G. França and José Bento. Markov Chain Lifting and Distributed ADMM. *IEEE Signal Processing Letters*, 24:294–298, 2017.

[2] 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*, 3(1):1–122, 2010.

[3] F. Chen, L. Lovász, and L. Pak. Lifting Markov Chains to Speed up Mixing. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 275–281, 1999.

[4] T. Erseghe, D. Zennaro, E. Dall'Anese, and L. Vangelista. Fast Consensus by the Alternating Direction Multipliers Method. *IEEE Transactions on Signal Processing*, 59(11):5523–5537, 2011.

[5] E. Ghadimi, A. Teixeira, M. G. Rabbat, and M. Johansson. The admm algorithm for distributed averaging: Convergence rates and optimal parameter selection. In *2014 48th Asilomar Conference on Signals, Systems and Computers*, pages 783–787, Nov 2014.

[6] X. Zhu, Z. Ghahramani, and J. Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *International Conference on Machine learning*, pages 912–919, 2003.

[7] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin. On the Linear Convergence of the ADMM in Decentralized Consensus Optimization. *IEEE Transactions on Signal Processing*, 62(7):1750–1761, 2014.

[8] Pontus Giselsson and Stephen Boyd. Diagonal scaling in douglas-rachford splitting and admm. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 5033–5039. IEEE, 2014.

[9] Pontus Giselsson and Stephen Boyd. Linear convergence and metric selection for douglas-rachford splitting and admm. *IEEE Transactions on Automatic Control*, 62(2):532–544, 2017.

[10] Wei Deng and Wotao Yin. On the global and linear convergence of the generalized alternating direction method of multipliers. *Journal of Scientific Computing*, 66(3):889–916, 2016.

[11] Mingyi Hong and Zhi-Quan Luo. On the Linear Convergence of the Alternating Direction Method of Multipliers. *Mathematical Programming*, 162(1):165–199, 2017.

[12] D. Davis and W. Yin. Convergence Rate Analysis of Several Splitting Schemes. arXiv:1406.4834, 2014.

[13] Damek Davis and Wotao Yin. Faster convergence rates of relaxed peaceman-rachford and admm under regularity assumptions. *Mathematics of Operations Research*, 2017.

[14] E. Wei and A. Ozdaglar. Distributed Alternating Direction Method of Multipliers. In *51st IEEE Conference on Decision and Control*, pages 5445–5450, 2012.

[15] Qing Ling, Wei Shi, Gang Wu, and Alejandro Ribeiro. Dlm: Decentralized linearized alternating direction method of multipliers. *IEEE Transactions on Signal Processing*, 63(15):4051–4064, 2015.

[16] Ali Makhdoumi and Asuman Ozdaglar. Broadcast-based distributed alternating direction method of multipliers. In *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, pages 270–277. IEEE, 2014.

[17] Ali Makhdoumi and Asuman Ozdaglar. Convergence rate of distributed admm over networks. *IEEE Transactions on Automatic Control*, 2017.

[18] Qing Ling, Yaohua Liu, Wei Shi, and Zhi Tian. Communication-efficient weighted admm for decentralized network optimization. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 4821–4825. IEEE, 2016.

[19] André Teixeira, Euhanna Ghadimi, Iman Shames, Henrik Sandberg, and Mikael Johansson. Optimal Scaling of the ADMM Algorithm for Distributed Quadratic Programming. In *52nd IEEE Conference on Decision and Control*, pages 6868–6873. IEEE, 2013.

[20] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal Parameter Selection for the Alternating Direction Method of Multipliers (ADMM): Quadratic Problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015.

[21] Arvind U Raghunathan and Stefano Di Cairano. Optimal step-size selection in alternating direction method of multipliers for convex quadratic programs and model predictive control. In *Proceedings of Symposium on Mathematical Theory of Networks and Systems*, pages 807–814, 2014.

[22] Arvind U Raghunathan and Stefano Di Cairano. Admm for convex quadratic programs: Q-linear convergence and infeasibility detection. *arXiv preprint arXiv:1411.7288*, 2014.

[23] Franck Iutzeler, Pascal Bianchi, Ph Ciblat, and Walid Hachem. Linear convergence rate for distributed optimization with the alternating direction method of multipliers. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 5046–5051. IEEE, 2014.

[24] Aryan Mokhtari, Wei Shi, Qing Ling, and Alejandro Ribeiro. Decentralized quadratically approximated alternating direction method of multipliers. In *Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on*, pages 795–799. IEEE, 2015.

[25] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem. Explicit Convergence Rate of a Distributed Alternating Direction Method of Multipliers. *IEEE Transactions on Automatic Control*, 61(4):892–904, 2016.

[26] G. França and J. Bento. An Explicit Rate Bound for Over-Relaxed ADMM. In *IEEE International Symposium on Information Theory, ISIT 2016, Barcelona, Spain, July 10-15*, pages 2104–2108, 2016.

[27] Daniel Boley. Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. *SIAM Journal on Optimization*, 23(4):2183–2207, 2013.

[28] Deren Han and Xiaoming Yuan. Local linear convergence of the alternating direction method of multipliers for quadratic programs. *SIAM Journal on numerical analysis*, 51(6):3446–3457, 2013.

[29] K. Jung, D. Shah, and J. Shin. Fast Gossip Through Lifted Markov Chains. In *Proc. Allerton Conf. on Comm., Control, and Computing, Urbana-Champaign, IL*, 2007.

[30] W. Li, H. Dai, and Y. Zhang. Location-Aided Fast Distributed Consensus in Wireless Networks. *IEEE Transactions on Information Theory*, 56(12):6208–6227, 2010.

[31] K. Jung, D. Shah, and J. Shin. Distributed Averaging via Lifted Markov Chains. *IEEE Transactions on Information Theory*, 56(1):634–647, 2010.

[32] J. R. Norris. *Markov Chains*. Cambridge University Press, 1998.

[33] Philipp Zumstein. Comparison of spectral methods through the adjacency matrix and the laplacian of a graph. *TH Diploma, ETH Zürich*, 2005.