21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Robustness of large-scale stochastic matrices to localized perturbations

Giacomo Como Fabio Fagnani

How much can the invariant probability distribution π of an irreducible row-stochastic matrix P be affected by perturbations localized on a relatively small subset $\mathcal{W}$ of its state space $\mathcal{V}$? Such a question arises in an increasing number of applications, most notably in the emerging field of large-scale networks.

As an example, many notions of network centrality can be formulated in terms of invariant probability distributions of suitably defined stochastic matrices. In particular, Google's PageRank algorithm [6] assigns to webpages values corresponding to the entries of the invariant probability distribution π of the matrix P obtained as a convex combination of the normalized adjacency matrix of the directed graph describing the hyperlink structure of the World Wide Web (WWW), and of a matrix whose all entries equal the inverse of the total number of webpages [22], [10]. A well-known problem in this context is rank-manipulation, i.e., the intentional addition or removal of hyperlinks from some webpages (hence, the alteration of the corresponding rows of P) with the goal of modifying the PageRank vector [4], [21], [13]. A natural question is then, to what extent a small subset $\mathcal{W}$ of webpages can alter the PageRank vector π. Similar robustness issues have been raised for accidental variations of the WWW topology occurring, e.g., because of server failures or network congestion problems [19].

More generally, the problem is of central interest in the context of distributed averaging and consensus algorithms [34]. There, linear systems of the form $x(t+1) = Px(t)$, or their continuous-time analogous, are studied, e.g., as algorithms for distributed optimization [39], [40], [5], control [20], [33], [7], synchronization in sensor networks [35], or reputation management in ad-hoc networks [26], as well as behavioral models for flocking phenomena [41], [14], or opinion dynamics in social networks [15], [16], [18], [1]. Equilibria of such systems are consensus vectors, i.e., multiples of the all-one vector, and standard results following from Perron-Frobenious theory guarantee convergence (with the additional assumption of aperiodicity of P, in the discrete time case) to a consensus vector with all entries equal to $\pi' x(0)$. Depending on the specific applicative context, the natural question is to what extent the consensus value $\pi' x(0)$ is affected by perturbations of P corresponding, e.g., to malfunctioning of a small fraction of the sensors, or conservative/influential minorities in social networks [2].

Other applications can be found in the context of interacting particle systems [24], [25]. In particular, in the voter model on a finite graph [11], [12], [3, Ch. 14], [17, Ch. 6.9], the probability distribution of the final consensus value is determined by the invariant distribution of the stochastic matrix associated to the simple random walk on the graph. Perturbations in this case may model the presence of inhomogeneities or 'zealots' [30], [31].

The above-described problems all boil down to estimating the distance between the invariant probability distribution π of an irreducible stochastic matrix P and an invariant probability distribution $\tilde{\pi} = \tilde{P}'\tilde{\pi}$ of another stochastic matrix $\tilde{P}$, to be interpreted as a perturbed version of P. In some applications, P may be reversible, i.e., coincide with the normalization of a symmetric positive matrix, so that π can be easily computed in terms of the latter. However, even in these cases, the considered perturbations will typically be such that $\tilde{P}$ is not reversible and thus $\tilde{\pi}$ does not allow for a tractable explicit expression.

Remarkably, standard perturbation results based on sensitivity analysis [36], [37], [38], [27], [8], [9], [28], [29], [2] do not provide a satisfactory answer to this problem. Indeed, they provide upper bounds of the form

$$||\tilde{\pi} - \pi||_p \leq \kappa_P ||\tilde{P} - P||_q \,, \tag{1}$$

for some $p, q \in [1, \infty]$, where κ_P is a condition number depending on the original stochastic matrix

Department of Automatic Control, Lund University, Box 118, SE-221 00, Lund, Sweden, giacomo.como@control.lth.se
Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, fabio.fagnani@polito.it

ISBN: 978-90-367-6321-9

P only. Such condition numbers are lower bounded by an absolute positive constant (e.g., $1/4$ for the smallest of those surveyed in [9]) and typically blow up as the state space $\mathcal{V}$ grows large. Therefore, such results do not allow one to prove that the distance $||\tilde{\pi} - \pi||_p$ vanishes in the limit of large network size, even if P and $\tilde{P}$ differ only in a single row, unless $||\tilde{P} - P||_q$ itself vanishes.

In this paper, we obtain upper bounds on the total variation distance $||\tilde{\pi} - \pi|| := \frac{1}{2}||\tilde{\pi} - \pi||_1$ of the form

$$||\tilde{\pi} - \pi|| \leq \theta\left(\tau \frac{\tilde{\chi}}{\tau^*_{\mathcal{W}}}\right), \qquad (2)$$

where $\theta : [0, +\infty) \to [0, 1]$

$$\theta(x) := \begin{cases} x \ln\left(e^2/x\right) & x \leq x^* \\ 1 & x \geq x^*, \end{cases} \qquad (3)$$

where $x^* = 0.31784\ldots$ is the smallest positive solution of $e^2/x = \exp(1/x)$, is a continuous, non-decreasing function such that $\theta(0) = 0$;

$$\tau := \inf\left\{t \geq 1 : ||P^t_{u,\cdot} - P^t_{v,\cdot}|| \leq 1/e\,, \forall u, v \in \mathcal{V}\right\} \qquad (4)$$

is the mixing time of the original stochastic matrix P; $\tau^*_{\mathcal{W}}$ denotes the minimal expected hitting time on the set $\mathcal{W}$ for a Markov chain with transition probability matrix P; and $\tilde{\chi}$ stands for the escape time from $\mathcal{W}$ for a Markov chain with transition probability matrix $\tilde{P}$. As opposed to the aforementioned sensitivity results, all derived from algebraic arguments, our proofs rely on coupling techniques, combined with an argument similar to the one developed in [1] for 'highly fluid' networks. Clearly, (2) implies that $||\tilde{\pi} - \pi||$ vanishes provided that $\tau\tilde{\chi}/\tau^*_{\mathcal{W}}$ does. As we will show, this finds immediate application in the PageRank manipulation problem. More in general, our results prove useful in many of those aforementioned large-scale network applications where classical sensitivity-based results fail to provide a satisfactory answer.

Mixing properties of stochastic matrices have been the object of extensive recent research [3], [32], [23], and several results are available allowing one to estimate the mixing time τ of a stochastic matrix P, e.g., in terms of the conductance or other geometrical properties of the graph associated to P. It is worth pointing out that a connection between mixing properties and robustness of stochastic matrices is already unveiled by the perturbation results of [28], [29], where (1) is proven for $p = 1$, $q = \infty$, and condition number κ_P proportional to τ. Of a similar flavor are Seneta's results [37], [38] estimating the condition number κ_P in terms of ergodicity coefficients. Also the estimates proposed in [2] for symmetric P can be rewritten as (1) with for $p = q = 2$ and κ_P equal to the inverse of the spectral gap of P. As compared to these references, the fundamental novelty of our bound (2) consists in measuring the size of the perturbation in terms of the ratio $\tilde{\chi}/\tau^*_{\mathcal{W}}$ instead of the distance $||\tilde{P} - P||_q$, thus enabling one to obtain significant results in scenarios where $\mathcal{W}$ is small but $\tilde{P} - P$ is not necessarily small in any norm.

In fact, of the last two parameters appearing in the righthand side of (2), the escape time $\tilde{\chi}$ is the only one truly depending on the perturbation $\tilde{P} - P$, and is indeed easily estimated in typical cases when $\mathcal{W}$ is a small subset of $\mathcal{V}$. On the other hand, the minimal hitting time $\tau^*_{\mathcal{W}}$, which depends on P and $\mathcal{W}$ only, turns out to be the hardest to get lower bounds on in typical applications where P is sparse and $\mathcal{W}$ remains small but not necessary localized as the state space grows large. While Kac's formula ([23, Lemma 21.13]) readily implies the upper bound $\tau^*_{\mathcal{W}} \leq 1/\pi(\mathcal{W})$, lower bounds on $\tau^*_{\mathcal{W}}$ typically involve finer details of P than just $\pi(\mathcal{W})$. We will propose an analysis of $\tau^*_{\mathcal{W}}$ for networks with high local connectivity, which finds natural application when the graph associated to P is a d-dimensional grid, and the size of $\mathcal{W}$ remains bounded (or grows very slowly) as the network size grows large. We will also discuss results for random, locally tree-like networks will be the object of a forthcoming work.

Note: This extended abstract is based on a recent journal submission of the authors available online at http://arxiv.org/abs/1309.0158. Please refer to it for full proofs and examples of the results discussed above.

References

[1] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, *Opinion fluctuations and disagreement in social networks*, Mathematics of Operation Research **38** (2013), no. 1, 1–27.

[2] D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi, *Spread of (mis)information in social networks*, Games and Economic Behavior **70** (2010), no. 2, 194–227.

[3] D. Aldous and J. Fill, *Reversible Markov chains and random walks on graphs*, Monograph in preparation, http://www.stat.berkeley.edu/ aldous/RWG/book.html, 2002.

[4] K. Avrachenkov and N. Litvak, *The effect of new links on Google PageRank*, Stochastic Models **22** (2006), 319–331.

[5] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and distributed computation: Numerical methods*, Athena Scientific, Belmont, MA, 1997.

[6] S. Brin and L. Page, *The anatomy of a large-scale hypertextual web search engine*, Proceedings of the 7th International World Wide Web Conference, 1998, pp. 107–117.

[7] R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri, *Communication constraints in coordinated consensus problems*, Proceedings of IEEE ACC, 2006.

[8] G.E. Cho and C.D. Meyer, *Markov chain sensitivity measured by mean first passage times*, Linear Algebra and its Applications **316** (2000), 21–28.
[9] ———, *Comparison of perturbation bounds for the stationary distribution of a Markov chain*, Linear Algebra and its Applications **335** (2001), 137–150.
[10] F. Chung, *The heat kernel as the pagerank of a graph*, Proceedings of the National Academy of Science **104** (2007), no. 50, 19735–19740.
[11] P. Connely and D. Welsh, *Finite particle systems and infection models*, Mathematical Proceedings of the Cambridge Philosophical Society **94** (1982), 167–182.
[12] J.T. Cox, *Coalescing random walks and voter model consensus times on the torus in* $\mathbb{Z}^d$, The Annals of Probability **17** (1989), no. 4, 1333–1366.
[13] B.C. Csáji, R.M. Jungers, and V.D. Blondel, *PageRank optimization by edge selection*, Applied Discrete Mathematics (2013).
[14] F. Cucker and S. Smale, *The mathematics of emergence*, Japanese Journal of Mathematics **2** (2007), no. 1, 197–227.
[15] M.H. DeGroot, *Reaching a consensus*, Journal of the American Statistical Association **69** (1974), no. 345, 118–121.
[16] P.M. DeMarzo, D. Vayanos, and J. Zwiebel, *Persuasion bias, social influence, and unidimensional opinions*, The Quarterly Journal of Economics **118** (2003), no. 3, 909–968.
[17] R. Durrett, *Random graph dynamics*, Cambridge University Press, 2006.
[18] B. Golub and M.O. Jackson, *Naive learning in social networks and the wisdom of crowds*, American Economic Journal: Microeconomics **2** (2010), no. 1, 112–149.
[19] H. Ishii and R. Tempo, *Computing the pagerank variation for fragile web data*, SICE Journal of Control, Measurement, and System Integration **2** (2006), no. 1, 1–9.
[20] A. Jadbabaie, J. Lin, and S. Morse, *Coordination of groups of mobile autonomous agents using nearest neighbor rules*, IEEE Transactions on Automatic Control **48** (2003), no. 6, 988–1001.
[21] C. De Kerchove, L. Ninove, and P. Van Dooren, *Maximizing PageRank via outlinks*, Linear Algebra and its Applications **429** (2008), 1254–1276.
[22] A.N. Langville and C.D. Meyer, *Google's pagerank and beyond: The science of search engine rankings*, Princeton University Press, 2006.
[23] D.A. Levin, Y. Peres, and E.L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, 2010.
[24] T.M. Liggett, *Interacting particle systems*, Springer-Verlag, 1985.
[25] ———, *Stochastic interacting systems: Contact, voter, and exclusion processes*, Springer Berlin, 1999.
[26] Y. Liu and Y.R. Yang, *Reputation propagation and agreement in mobile ad hoc networks*, Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), 2003, pp. 1510–1515.
[27] C.D. Meyer, *Sensitivity of the stationary distribution of a Markov chain*, SIAM Journal on Matrix Analysis and Applications **15** (1994), 715–728.
[28] A. Mitrophanov, *Stability and exponential convergence of continuous-time Markov chains*, Journal of Applied Probability **40** (2003), 970–979.
[29] ———, *Sensitivity and convergence of uniformly ergodic Markov chains*, Journal of Applied Probability **42** (2005), 1003–1014.
[30] M. Mobilia, *Does a single zealot affect an infinite group of voters?*, Physical Review Letters **91** (2003), no. 2, 028701.
[31] M. Mobilia and I.T. Georgiev, *Voting and catalytic processes with inhomogeneities*, Physical Review E **71** (2005), 046102.
[32] R. Montenegro and P. Tetali, *Mathematical aspects of mixing times in Markov chains*, Foundations and Trends in Theoretical Computer Science **1** (2006), no. 3, 237–354.
[33] R. Olfati-Saber and R.M. Murray, *Consensus problems in networks of agents with switching topology and time-delays*, IEEE Transactions on Automatic Control **49** (2004), no. 9, 1520–1533.
[34] A. Olshevsky and J.N. Tsitsiklis, *Convergence speed in distributed consensus and averaging*, SIAM Review **53** (2011), no. 4, 747–772.
[35] L. Qun and D. Rus, *Global clock synchronization in sensor networks*, IEEE Journal on Computers **55** (2006), no. 2, 214–226.
[36] P. J. Schweitzer, *Perturbation theory and finite Markov chains*, Journal of Applied Probability **5** (1968), 401–413.
[37] E. Seneta, *Perturbation of the stationary distribution measured by ergodicity constraints*, Advances in Applied Probability **20** (1988), 228–230.
[38] ———, *Sensitivity of finite Markov chains under perturbation*, Statistics & Probability Letters **17** (1993), 163–168.
[39] J.N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1984.
[40] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans, *Distributed asynchronous deterministic and stochastic gradient optimization algorithms*, IEEE Transactions on Automatic Control **31** (1986), no. 9, 803–812.
[41] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and Schochet O., *Novel type of phase transitions in a system of self-driven particles*, Physical Review Letters **75** (1995), no. 6, 1226–1229.