# SMALL PERTURBATION ANALYSIS OF NETWORK TOPOLOGIES

*Elena Ceci and Sergio Barbarossa*

Dept. of Information Engineering, Electronics, and Telecommunications, Sapienza University of Rome,
Via Eudossiana,18, 00184, Rome, Italy

## ABSTRACT

The goal of this paper is to derive a small perturbation analysis for networks subject to random changes of a small number of edges. Small perturbation theory allows us to derive, albeit approximate, closed form expressions that make possible the theoretical statistical characterization of the network topology changes. The analysis is instrumental to formulate a graph-based optimization algorithm, which is robust against edge failures. In particular, we focus on the optimal allocation of the overall transmit powers in wireless communication networks subject to fading, aimed at minimizing the variation of the network connectivity, subject to a constraint on the overall power necessary to maintain network connectivity.

***Index Terms—*** *Graph topologies*, *network reliability*, *information flow*, *small perturbation analysis*

## 1. INTRODUCTION

Graph theory is a powerful mathematical tool to extract macroscopic properties of interconnected entities. In some cases, like sensor, communication and transportation networks, there exists a physical link (edge) between pairs of vertices. In other cases, the data are represented as points in a high-dimensional space and edges are associated to pairs of points to reflect their similarity. These representations form the basis of many graph-based unsupervised or semi-supervised machine learning techniques. In several cases, the presence of an edge between a pair of nodes is subject to random changes. In a wireless communication system, for instance, it is typical to have random link failures due to fading. Similarly, in a point cloud, the association of an edge to a pair of points can also be a random event, because of imperfect information in the rule used to decide whether to establish an edge or not. The goal of this paper is to assess the effect of random changes on a limited number of edges on macroscopic network parameters, such as, for example, connectivity. We build our study on a small perturbation analysis of the eigendecomposition of the Laplacian matrix describing the graph. An outcome of our analysis is the identification of the most critical links, i.e. those links whose failure has a major impact on some network macroscopic features, such as connectivity. A nearby research field is network reliability, [1], [2], [3], and [4], assessing the sensitivity of the network to random link failures. In particular, if a failure probability is associated to every edge, the all-terminal reliability is the probability that the entire graph is connected, while the $K$-terminal probability is the probability that subset of $K$ nodes are connected by one or more paths [1]. In [4], for example, it was proposed a method to solve a network reliability optimization problem, where reliability and cost are, respectively, the objective function to be maximized and the constraint. Eubank et al. in [2] focused on the infectious disease outbreaks over complex networks. In such a case, the reliability is used to characterize the criticalities of the network and to identify the set of edge or vertices whose deletions mostly affects the disease diffusion. In our work, we also associate a parameter of importance (criticality) to every edge. This parameter is expressed in closed form by using a small perturbation analysis [5] of the Laplacian matrix eigen-decomposition, valid when the percentage of perturbed (either added or deleted) edges is small. There are recent works that apply matrix perturbation theory to spectral clustering, such as [6], [7], [8], [9], and [10]. In particular, Von Luxburg in [6] used perturbation theory to analyze clustering, evaluating the distance between the eigenspace spanned by sets of Laplacian eigenvectors of a nominal graph and the corresponding eigenspace of a perturbed Laplacian. Spielman in [11] recalled the basics of matrix perturbation theory to analyze spectral partitioning heuristics on random graphs that are generated to have good partitions.

The goal of our paper is threefold: i) we introduce a new measure of edge centrality based on perturbation analysis; ii) we provide a statistical analysis of network macroscopic parameters such as connectivity, due to random link failure using a small perturbation analysis of the Laplacian eigen-decomposition; iii) we provide an optimal wireless network resource allocation where we minimize the expected perturbation of the algebraic connectivity subject to a constraint on the overall cost (power) needed to maintaining the links, in the presence of fading.

This work has been supported by H2020 EUJ Project 5G-MiEdge, Nr. 723171.

## 2. STATISTICAL ANALYSIS OF CONNECTIVITY BASED ON SMALL PERTURBATION

In this section, we briefly introduce the mathematical tools of graph-based methods and we provide a small perturbation analysis of the Laplacian eigendecomposition. The analysis is valid in cases where only a few percentage of edges is perturbed. Then, we derive a statistical analysis of the algebraic connectivity of a network subject to random edge perturbation. A small perturbation analysis of the eigen-decomposition of a matrix is a classical problem that has been studied for a long time, see, e.g. [5], [12]. In this paper we focus on the small perturbation analysis of the eigendecomposition of a perturbed Laplacian $\boldsymbol{L} + \delta \boldsymbol{L}$, incorporating an original graph Laplacian $\boldsymbol{L}$ plus the addition or deletion of a small percentage of edges. We denote by $\tilde{\lambda}_i = \lambda_i + \Delta\lambda_i$ the perturbed $i$-th eigenvalue and by $\tilde{\boldsymbol{u}}_i = \boldsymbol{u}_i + \Delta\boldsymbol{u}_i$ the associated perturbed eigenvector. If only one edge fails, say edge $m$, the perturbation matrix can be written as $\delta\boldsymbol{L}(m) = -\boldsymbol{a}_m\boldsymbol{a}_m^T$, where $\boldsymbol{a}_m = [a_{m_1} \cdots a_{m_n}]^T$ has all entries equal to zero, except the two $a_m(i_m) = 1$ and $a_m(f_m) = -1$, where $i_m$ and $f_m$ are the initial and final vertices of the failing edge. In case of addition of a new edge, the perturbation matrix is simply the opposite of the previous expression, i.e. $\delta\boldsymbol{L}(m) = \boldsymbol{a}_m\boldsymbol{a}_m^T$. It is straightforward to see that the perturbation of the Laplacian matrix due to the simultaneous deletion of a small set of edges is simply $\delta\boldsymbol{L} = -\sum_{m\in\mathcal{E}_p}\boldsymbol{a}_m\boldsymbol{a}_m^T$ where $\mathcal{E}_p$ denotes the set of perturbed edges. In the case where all eigenvalues are distinct and the perturbation affects a few percentage of links, the perturbed eigenvalues and eigenvectors $\tilde{\lambda}_i$ and $\tilde{\boldsymbol{u}}_i$ are related to the unperturbed values $\lambda_i$ and $\boldsymbol{u}_i$ by the following formulas [5]:

$$\tilde{\lambda}_i \simeq \lambda_i + \boldsymbol{u}_i^T \delta\boldsymbol{L}\,\boldsymbol{u}_i \tag{1}$$

$$\tilde{\boldsymbol{u}}_i \simeq \boldsymbol{u}_i + \sum_{j\neq i} \frac{\boldsymbol{u}_j^T \delta\boldsymbol{L}\,\boldsymbol{u}_i}{\lambda_i - \lambda_j}\boldsymbol{u}_j. \tag{2}$$

Hence, the perturbations $\Delta\lambda_i(m)$ of the $i$-th eigenvalue and that of the associated eigenvector $\Delta\boldsymbol{u}_i(m)$, caused by the deletion of edge $m \in \mathcal{E}$, are:

$$\begin{aligned}\Delta\lambda_i(m) &= \boldsymbol{u}_i^T\delta\boldsymbol{L}(m)\boldsymbol{u}_i = -\boldsymbol{u}_i^T\boldsymbol{a}_m\boldsymbol{a}_m^T\boldsymbol{u}_i \\ &= -||\boldsymbol{a}_m^T\boldsymbol{u}_i||^2 = -[u_i(f_m) - u_i(i_m)]^2;\end{aligned} \tag{3}$$

$$\begin{aligned}\Delta\boldsymbol{u}_i(m) &= \sum_{j\neq i}\frac{\boldsymbol{u}_j^T\delta\boldsymbol{L}(m)\boldsymbol{u}_i}{\lambda_i-\lambda_j}\boldsymbol{u}_j = -\sum_{j\neq i}\frac{\boldsymbol{u}_j^T\boldsymbol{a}_m\boldsymbol{a}_m^T\boldsymbol{u}_i}{\lambda_i-\lambda_j}\boldsymbol{u}_j \\ &= \sum_{j\neq i}\frac{[u_j(i_m)-u_j(f_m)][u_i(f_m)-u_i(i_m)]}{\lambda_i-\lambda_j}\boldsymbol{u}_j.\end{aligned} \tag{4}$$

Within the limits of validity of first order perturbation analysis, the overall perturbation resulting from the deletion of multiple edges in a set $\mathcal{E}_p$ of perturbed edges is the sum of all the perturbations occurring on single edges:

$$\Delta\lambda_i = \sum_{m\in\mathcal{E}_p}\Delta\lambda_i(m). \tag{5}$$

In their simplicity, the above formulas capture some of the most relevant aspects of perturbation and their relation to graph topology. In fact, it is known from spectral graph theory, see e.g., [6], that the entries of the Laplacian eigenvectors associated to the smallest eigenvalues tend to be smooth and assume the same sign over vertices within a cluster, while they can vary arbitrarily across different clusters. Taking into account these properties, the above perturbation formulas (1)-(4) can be interpreted as follows:

1. the edges whose deletion causes the largest perturbation are inter-cluster edges;
2. given a connected graph, the eigenvector associated to the null eigenvalue does not induce any perturbation on any other eigenvalue/eigenvector, because it is constant;
3. eigenvectors associated to eigenvalues very close to nearby eigenvalues typically suffer larger perturbations (recall that formulas (1) and (2) hold true only for distinct eigenvalues).

We provide now a simple statistical analysis of the eigenvalue perturbation, valid in case of deletion of a generic edge $m$, which is modeled as a random event occurring with probability $\mathbb{P}(m)$. Denoting by $Z_m$ a binary random variable that assumes value 0, with probability $1-\mathbb{P}(m)$, or 1, with probability $\mathbb{P}(m)$, the absolute value of the perturbation of eigenvalue $\lambda_i$ is approximately (within the validity of small perturbation theory):

$$Y^{(i)} := |\Delta\lambda_i| = \sum_{m=1}^{M} Z_m|\Delta\lambda_i(m)|, \tag{6}$$

with $\Delta\lambda_i(m)$ defined as in (3). Since we use small perturbation analysis, the validity of (6) holds when the probabilities $\mathbb{P}(m)$ are sufficiently small. How small, it depends on the number of edges in the graph. As a rule of thumb, if $M$ is the number of edges, $\mathbb{P}(m)$ should not be much larger than $1/M$, so that the number of perturbed edges is small. Assuming that the events associated to the deletion of different edges are statistically independent, it is easy to derive the statistical properties of the eigenvalues' perturbations. In particular, expected value and variance of $Y^{(i)}$ are:

$$m_{Y^{(i)}} := \mathbb{E}\{Y^{(i)}\} = \sum_{m=1}^{M}\mathbb{P}_m|\Delta\lambda_i(m)| \tag{7}$$

and

$$\mathrm{var}\{Y^{(i)}\} = \sum_{m=1}^{M}(1-\mathbb{P}_m)\mathbb{P}_m|\Delta\lambda_i(m)|^2 \tag{8}$$

Furthermore, we can derive a bound on the probability that the random variable $Y^{(i)}$ does not deviate from its expected value more than a given value $t$. This probability can be upper bounded using Hoeffding's bound [13], which enables us to write

$$\mathbb{P}\{Y^{(i)} - m_{Y^{(i)}} \geq t\} \leq e^{-2t^2/\sum_{m=1}^{M} |\Delta\lambda_i(m)|^2}. \quad (9)$$

In particular, the probability that the eigenvalue perturbation be larger than a certain percentage $\alpha \in [0, 1]$ of the true eigenvalue $\lambda_i$, is upper bounded by

$$\mathbb{P}\{Y^{(i)} \geq \alpha\lambda_i\} = \mathbb{P}\{Y^{(i)} - m_{Y^{(i)}} \geq \alpha\lambda_i - m_{Y^{(i)}}\} \leq e^{-2(\alpha\lambda_i - m_{Y^{(i)}})^2/\sum_{m=1}^{M} |\Delta\lambda_i(m)|^2}. \quad (10)$$

## 3. A NEW MEASURE OF EDGE CENTRALITY

Based on the above derivations, we propose a new measure of edge centrality, which we call *perturbation centrality*. If the cluster is composed of $K$ clusters, we define the topology perturbation centrality of edge $m$ as follows:

$$\mathsf{p}_K(m) := \sum_{i=2}^{K} |\Delta\lambda_i(m)| \quad (11)$$

where $\Delta\lambda_i(m)$ is defined as in (3). The summation starts from $i = 2$ simply because, from (1), the perturbation induced by the deletion of any edge on the smallest eigenvalue is null. The above parameter $\mathsf{p}_K(m)$ assigns to each edge the perturbation that its deletion causes to the overall network connectivity, measured as the sum of the $K$ smallest eigenvalues of the Laplacian matrix [6]. This parameter is particularly relevant in case of modular graphs, i.e. graphs evidencing the presence of clusters. In such a case, it is well known from spectral clustering theory [6] that the smallest eigenvalues of the Laplacian carry information about the number of clusters in a graph. In Fig.1 we report an example of modular graph, obtained by connecting two clusters through a few edges. The *perturbation centrality* is encoded in the color intensity of each edge. It is interesting to see that the edges with the darkest color are, as expected, the inter-cluster edges.

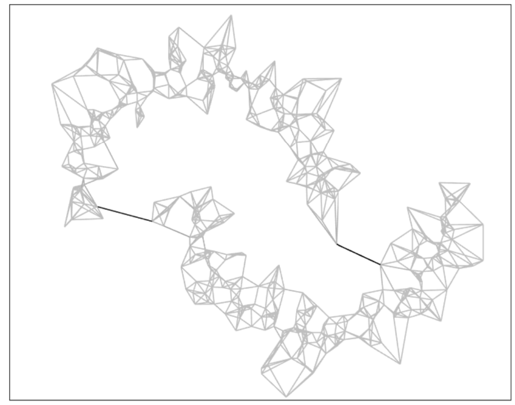

**Fig. 1**: Example of *perturbation centrality measure* .

## 4. APPLICATION: ROBUST INFORMATION TRANSMISSION OVER WIRELESS NETWORKS

Now we apply our statistical analysis to optimize the resource (power) allocation over a wireless network in order to make the network robust against random link failures. We consider a wireless communication network where the links are subject to random failure because of fading. Every edge is characterized by an outage probability $P_{out}(m), m = 1, \ldots, M$. We suppose the failure events over different links to be independent of each other. We consider first a single-input-single-output (SISO) Rayleigh flat fading channel for each link. In such a case, the channel coefficient $h$ is a complex Gaussian random variable (rv) with zero mean and circularly symmetric. Hence, the random variable $\alpha = |h|^2$ has an exponential distribution. Denoting by $F_n(x; \lambda)$ the cumulative distribution function (CDF) of a gamma random variable of order $n$, with parameter $\lambda$, the CDF of $\alpha$ can then be written as $F_1(\alpha, \lambda)$. We also denote with $C = log_2(1 + |h|^2\rho)$ the link capacity (in bits/sec/Hz), where $\rho = \frac{P_T(m)}{\sigma_n^2 r_m^2}$ is the signal-to-noise ratio (SNR), $P_T(m)$ is the transmitted power over the $m$-th link, $\sigma_n^2$ is the noise variance, and $r_m$ the distance covered by link $m$. Denoting by $R$ the data rate, the outage probability $P_{out}(m)$ is defined as:

$$P_{out}(m) = Pr\{C < R\} = \quad (12)$$
$$Pr\{\log_2(1 + |h|^2\rho) < R\} = Pr\{|h|^2 < \frac{2^R - 1}{\rho}\} =$$
$$\int_0^{\frac{2^R-1}{\rho}} \lambda e^{-\lambda\alpha} d\alpha = F_1\left(\frac{2^R - 1}{\rho}, \lambda\right) = 1 - e^{-\frac{\lambda}{\rho}(2^R-1)}$$

This expression can be inverted to derive the transmit power $P_T(m)$ as a function of the outage probability:

$$P_T(m) = -\frac{\lambda\sigma_n^2 r_m^2 (2^R - 1)}{log(1 - P_{out}(m))} = \frac{\sigma_n^2 r_m^2 (2^R - 1)}{F_1^{-1}(P_out(m); \lambda)}. \quad (13)$$

The small perturbation statistical analysis derived in Section 2 is instrumental now to formulate a robust network optimization problem. We assess the network robustness, in terms of connectivity, as the ability of the network connectivity to be slightly affected by the failure of a small number of edges. The network connectivity is measured by the second smallest eigenvalue of the Laplacian, also known as the graph *algebraic connectivity*. This parameter is known to provide a bound for the graph conductance [14]. Our goal is to evaluate the transmit powers $P_T(m)$, or equivalently, through (13), the outage probabilities, that minimize the average perturbation of the algebraic connectivity, subject to a cost associated to the total transmit power $P_{T_{max}}$ necessary to establish all the links. In formulas, we wish to solve the following optimization problem:

$$\begin{aligned} \min_{\mathbf{P}_{out}} \quad & \sum_{m\in\mathcal{E}} \mathbb{E}\{|\Delta\lambda_2(m)|\} \\ s.t \quad & \sum_{m\in\mathcal{E}} P_T(m) \leq P_{T_{max}} \\ & P_{out}(m) \in [0,1], \forall m \in \mathcal{E}. \end{aligned} \tag{14}$$

Using equations (7) and (3), we can rewrite the optimization problem in terms of the outage probabilities $P_{out}(m)$ as:

$$\begin{aligned} \min_{\mathbf{P}_{out}} \quad & \sum_{m\in\mathcal{E}} P_{out}(m)[u_2(i_m) - u_2(f_m)]^2 \\ \text{s.t.} \quad & \sum_{m\in\mathcal{E}} \frac{r_m^2}{F_1^{-1}(P_{out}(m);\lambda)} \leq C_{max} \\ & P_{out}(m) \in [0,1], \; \forall\, m \in \mathcal{E} \end{aligned} \tag{15}$$

where we set $C_{max} = \frac{P_{T_{max}}}{\sigma_n^2(2^R-1)}$.

Problem (15) is non-convex because the constraint set is not convex. If we perform the change of variable $t_m := 1/F_1^{-1}(P_out(m);\lambda) = -\lambda/log(1-P_{out}(m)), m = 1,\ldots,M$, the constraint becomes linear and the objective function $F_1(\frac{1}{t_m};\lambda)|\Delta\lambda_2(m)| = \sum_{m\in\mathcal{E}}(1-e^{-\frac{\lambda}{t_m}})|\Delta\lambda_2(m)|$ becomes non-convex. However, if we limit the variability of the unknown variables to the set $t_m \geq \lambda/2, \forall m$, the objective function becomes convex, so that the original problem converts into the following convex problem:

$$\begin{aligned} \min_{\mathbf{t}} \quad & \sum_{m\in\mathcal{E}} (1-e^{-\frac{\lambda}{t_m}})|\Delta\lambda_2(m)| \\ s.t. \quad & \sum_{m\in\mathcal{E}} r_m^2 t_m \leq C_{max} \\ & t_m \geq \tfrac{\lambda}{2}, \quad \forall m \in \mathcal{E}. \end{aligned} \tag{16}$$

We can now generalize the previous formulation to the Multi-Input Multi-Output (MIMO) case, assuming multiple independent Rayleigh fading channels. A fundamental property of MIMO systems is the diversity gain, which makes them more robust against fading with respect to SISO systems. In a MIMO channel with $n_T$ transmit and $n_R$ receive antennas, if the $n = n_T \times n_R$ channels are statistically independent, denoting by $h_{ij}$ the coefficient between the $i$-th transmit and the $j$-th receive antenna, the pdf of the random variable $\alpha := \sum_{i=1}^{n_T}\sum_{j=1}^{n_R}|h_{ij}|^2$ is the Gamma distribution [15]:

$$P_A(\alpha) = \frac{\lambda^n}{(n-1)!}\alpha^{n-1}e^{-\lambda\alpha}. \tag{17}$$

Proceeding similarly to the SISO case, the optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{t}} \quad & \sum_{m\in\mathcal{E}} F_n(\tfrac{1}{t_m};\lambda)|\Delta\lambda_2(m)| \\ s.t. \quad & \sum_{m\in\mathcal{E}} r_m^2 t_m \leq C_{max} \\ & t_m \geq \lambda/(n+1), \quad \forall m \in \mathcal{E} \end{aligned} \tag{18}$$

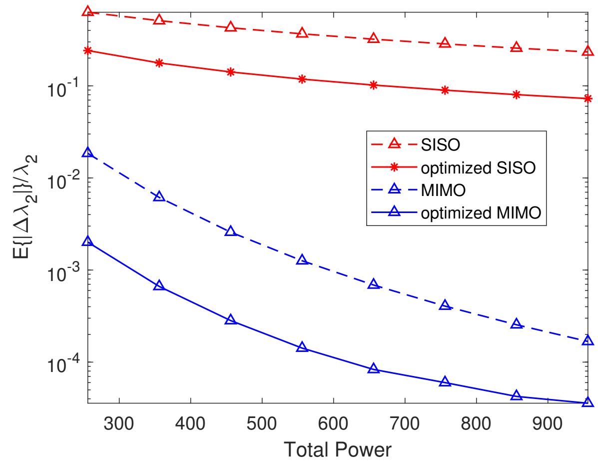


**Fig. 2**: Comparisons between SISO (red curves) and MIMO ($n = 4$) systems (blue curves), with and without optimization.

where the constraint on the variables $t_m$ has been introduced to make the problem convex. Indeed problem (16) is a special case of problem (18), when $n = 1$ and for $b = 1/\lambda$. An interesting result about the convexity of problem (18) is that the bounding region increases with the number of independent channels.

As a numerical example, we considered a connected network composed by two clusters, with a total of $|\mathcal{E}| = 1612$ edges and two bridge edges between the two clusters. For the sake of simplicity, we assumed the same distances $r_m$ over all links. In Fig. 2, we compare the expected perturbations of the algebraic connectivity, normalized to the nominal value $\lambda_2$, obtained using our optimization procedure or using the same power over all links, assuming the same overall power consumption. We report the result for both SISO and MIMO cases. From Fig. 2, we can observe a significant gain in terms of the total power necessary to achieve the same expected perturbation of the network algebraic connectivity. We can also see the advantage of using MIMO communications, at least in the case of statistically independent links.

## 5. CONCLUSIONS

In this paper we have shown how a small perturbation analysis of the graph Laplacian matrix can be beneficial to understand the effect of a few edge failures on the overall network connectivity. The closed form expressions resulting from this analysis enabled us to introduce a new edge centrality measure useful to assess how the failure of each edge affects the connectivity properties of the overall network. Furthermore, we used this analysis to formulate an optimal power allocation strategy over wireless communication networks, aimed to gain robustness against random link failures.

## 6. REFERENCES


[1] K. S. Tittmann P., *Graph-Based Modelling in Engineering*. Switzerland: Springer, Cham, 2017, ch. Structural Importance and Local Importance in Network Reliability.

[2] S. Eubank, M. Youssef, and Y. Khorramzadeh, "Using the network reliability polynomial to characterize and design networks," *Journal of complex networks 2.4*, p. 356372, September 2014.

[3] F. Beichelt and P. Tittmann, *Reliability and Maintenance: Networks and Systems*. Taylor & Francis, 2012.

[4] R. Dash, N. Barpanda, P. Tripathy, and C. Tripathy, "Network reliability optimization problem of interconnection network under node-edge failure model," *Applied Soft Computing*, vol. 12, no. 8, pp. 2322 – 2328, 2012.

[5] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. New York, NY, USA: Oxford University Press, Inc., 1988.

[6] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec 2007.

[7] S. Borjigin and C. Guo, "Perturbation analysis for the normalized laplacian matrices in the multiway spectral clustering method," *Science China Information Sciences*, vol. 57, no. 11, pp. 1–17, Nov 2014.

[8] Z. Tian, X. Li, and Y. Ju, "Spectral clustering based on matrix perturbation theory," *Science in China Series F: Information Sciences*, vol. 50, no. 1, pp. 63–81, 2007.

[9] L. Huang, D. Yan, N. Taft, and M. I. Jordan, "Spectral clustering with perturbed data," in *Advances in Neural Information Processing Systems*, 2009, pp. 705–712.

[10] A. J. Savol and C. S. Chennubhotla, "Approximating frustration scores in complex networks via perturbed laplacian spectra," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 92, no. 0, pp. 062 806–062 806, Dec 2015.

[11] D. A. Spielman, "Spectral graph theory and its applications," in *Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on*. IEEE, 2007, pp. 29–38.

[12] G. Stewart, *Introduction to matrix computations*, ser. Computer science and applied mathematics. Academic Press, 1973.

[13] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American statistical association*, vol. 58, no. 301, pp. 13–30, 1963.

[14] M. Newman, *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc., 2010.

[15] S. Barbarossa, *Multiantenna Wireless Communication Systems*, ser. Mobile Communications Series. Artech House, 2003.