Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications

Other titles in Foundations and Trends® in Machine Learning

Spectral Learning on Matrices and Tensors
Majid Janzamin, Rong Ge, Jean Kossaifi and Anima Anandkumar
ISBN: 978-1-68083-640-0

An Introduction to Variational Autoencoders
Diederik P. Kingma and Max Welling
ISBN: 978-1-68083-622-6

Elements of Sequential Monte Carlo
Christian A. Naesseth, Fredrik Lindsten and Thomas B. Schon
ISBN: 978-1-68083-632-5

Introduction to Multi-Armed Bandits
Aleksandrs Slivkins
ISBN: 978-1-68083-620-2

Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications

Ljubiša Stanković
University of Montenegro
Montenegro
ljubisa@ucg.ac.me

Danilo Mandic
Imperial College London
UK
d.mandic@imperial.ac.uk

Miloš Daković
University of Montenegro
Montenegro
milos@ucg.ac.me

Miloš Brajović
University of Montenegro
Montenegro
milosb@ucg.ac.me

Bruno Scalzo
Imperial College London
UK
bruno.scalzo-dees12@imperial.ac.uk

Shengxi Li
Imperial College London
UK
shengxi.li17@imperial.ac.uk

Anthony G. Constantinides
Imperial College London
UK
a.constantinides@imperial.ac.uk

the essence of knowledge
Boston — Delft

Foundations and Trends® in Machine Learning

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

L. Stanković, D. Mandic, M. Daković, M. Brajović, B. Scalzo, S. Li and A. G. Constantinides. *Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications*. Foundations and Trends® in Machine Learning, vol. 13, no. 4, pp. 332–530, 2020.

ISBN: 978-1-68083-980-7

Foundations and Trends® in Machine Learning

Volume 13, Issue 4, 2020

Editorial Scope

Topics

Foundations and Trends® in Machine Learning publishes survey and tutorial articles in the following topics:

- Adaptive control and signal processing
- Applications and case studies
- Behavioral, cognitive and neural learning
- Bayesian learning
- Classification and prediction
- Clustering
- Data mining
- Dimensionality reduction
- Evaluation
- Game theoretic learning
- Graphical models
- Independent component analysis
- Inductive logic programming
- Kernel methods
- Markov chain Monte Carlo
- Model choice
- Nonparametric methods
- Online learning
- Optimization
- Reinforcement learning
- Relational learning
- Robustness
- Spectral methods
- Statistical learning theory
- Variational inference
- Visualization

Information for Librarians

Foundations and Trends® in Machine Learning, 2020, Volume 13, 6 issues. ISSN paper version 1935-8237. ISSN online version 1935-8245. Also available as a combined paper and online subscription.

Contents

Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications

Ljubiša Stanković[1], Danilo Mandic[2], Miloš Daković[3], Miloš Brajović[4], Bruno Scalzo[5], Shengxi Li[6] and Anthony G. Constantinides[7]

[1] *University of Montenegro, Montenegro; ljubisa@ucg.ac.me*
[2] *Imperial College London, UK; d.mandic@imperial.ac.uk*
[3] *University of Montenegro, Montenegro; milos@ucg.ac.me*
[4] *University of Montenegro, Montenegro; milosb@ucg.ac.me*
[5] *Imperial College London, UK; bruno.scalzo-dees12@imperial.ac.uk*
[6] *Imperial College London, UK; shengxi.li17@imperial.ac.uk*
[7] *Imperial College London, UK; a.constantinides@imperial.ac.uk*

ABSTRACT

Modern data analytics applications on graphs often operate on domains where graph topology is not known a priori, and hence its determination becomes part of the problem definition, rather than serving as prior knowledge which aids the problem solution. Part III of this monograph starts by a comprehensive account of ways to learn the pertinent graph topology, ranging from the simplest case where the physics of the problem already suggest a possible graph structure, through to general cases where the graph structure is to be learned from the data observed on a graph. A particular emphasis is placed on the use of standard "relationship measures" in this context, including the correlation and precision matrices, together with the ways to combine these

Ljubiša Stanković, Danilo Mandic, Miloš Daković, Miloš Brajović, Bruno Scalzo, Shengxi Li and Anthony G. Constantinides (2020), "Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications", Foundations and Trends® in Machine Learning: Vol. 13, No. 4, pp 332–530. DOI: 10.1561/2200000078-3.

with the available prior knowledge and structural conditions, such as the smoothness of the graph signals or sparsity of graph connections. Next, for learning sparse graphs (that is, graphs with a small number of edges), the utility of the least absolute shrinkage and selection operator, known as (LASSO) is addressed, along with its graph specific variant, the graphical LASSO. For completeness, both variants of LASSO are derived in an intuitive way, starting from basic principles. An in-depth elaboration of the graph topology learning paradigm is provided through examples on physically well defined graphs, such as electric circuits, linear heat transfer, social and computer networks, and spring-mass systems. We also review main trends in graph neural networks (GNN) and graph convolutional networks (GCN) from the perspective of graph signal filtering. Particular insight is given to the role of diffusion processes over graphs, to show that GCNs can be understood from the graph diffusion perspective. Given the largely heuristic nature of the existing GCNs, their treatment through graph diffusion processes may also serve as a basis for new designs of GCNs. Tensor representation of lattice-structured graphs is next considered, and it is shown that tensors (multidimensional data arrays) can be treated as a special class of graph signals, whereby the graph vertices reside on a high-dimensional regular lattice structure. Finally, the concept of graph tensor networks is shown to provide a unifying framework for learning of big data on irregular domains. This part of monograph concludes with an in-dept account of emerging applications in financial data processing and underground transportation network modeling. More specifically, by means of portfolio cuts of an asset graph, we show how domain knowledge can be meaningfully incorporated into investment analysis, while the underground transportation example addresses vulnerability of stations in the London underground network to traffic disruption.

Keywords: graph theory; random data on graphs; big data on graphs; signal processing on graphs; machine learning on graphs; graph topology learning; systems on graphs; vertex-frequency estimation; graph neural networks; graphs and tensors.

1

Introduction

Graph data analytics have already shown enormous potential, as their flexibility in the choice of graph topologies (irregular data domains) and connections between the entities (vertices) allows for both a rigorous account of irregularly spaced information such as locations and social connections, and also for the incorporation of semantic and contextual cues, even for otherwise regular structures such as images.

In Part I and Part II of this monograph, it was assumed that the graph itself is already defined prior to analyzing data on graphs. The focus of Part I has been on defining graph properties through the mathematical formalism of linear algebra, while Part II introduces graph counterparts of several important standard data analytics algorithms, again for a given graph. However, in many modern applications, graph topology is not known a priori (Cioacă *et al.*, 2019; Das *et al.*, 2017; Dong *et al.*, 2015, 2016; Epskamp and Fried, 2018; Friedman *et al.*, 2008; Hamon *et al.*, 2019, Meinshausen *et al.*, 2006; Pavez and Ortega, 2016; Pourahmadi, 2011; Rabiei *et al.*, 2019; Stanković *et al.*, 2018, 2020), and the focus of this part is therefore on simultaneous estimation of data on a graph and the underlying graph topology. Without loss of generality, it is convenient to assume that the vertices are given, while the edges and their associated weights are part of the solution to the problem considered and need to be estimated from the vertex geometry and/or the observed data (Bohannon *et al.*, 2019; Caetano *et al.*, 2009; Camponogara and Nazari, 2015; Dal Col *et al.*, 2019; Gu and Wang, 2019; Mao and Gu, 2019; Pasdeloup *et al.*, 2019; Slawski and Hein, 2015; Segarra *et al.*, 2016; Stanković and Sejdić, 2019; Stanković *et al.*, 2017; Tanaka and Sakiyama, 2019; Thanou *et al.*, 2014; Ubaru *et al.*, 2017; Yankelevsky and Elad, 2016; Zhao *et al.*, 2012; Zheng *et al.*, 2011).

Three scenarios for the estimation of graph edges from vertex geometry or data are considered in this part of the monograph.

- Based on the *geometry of vertex positions.* In various sensor network setups (such as temperature, pressure, and transportation), the locations of the sensing positions (vertices) are known beforehand, while the vertex distances convey physical meaning about data dependence and thus may be employed for edge/weight determination.

- Based on *data association and data similarity.* Various statistical measures are available to serve as data association metrics, with the covariance and precision matrices most commonly used. A strong correlation between data on two vertices would indicate a large weight associated with the corresponding edge. A small degree of correlation would indicate nonexistence of an edge (after weight thresholding).

- *Based on physically well defined relations among the sensing positions.* Examples include electric circuits, power networks, linear heat transfer, social and computer networks, spring-mass systems, to mention but a few. In these cases, edge weighting can usually be well defined based on the underlying context of the considered problem.

After a detailed elaboration of graph definition and graph topology learning techniques, a summary of graph topology learning from data using probabilistic generative models is given. This followed by an account of graph neural networks (GNN), with a special emphasis on graph convolutional networks (GCN). The analysis is considered from the perspective of graph signal filtering presented in Part II. Graph data analysis is further generalized to the tensor representation of lattice-structured graphs, whereby the graph vertices reside on a high-dimensional tensor structure. Finally, two applications of graph-based data analysis are given: (i) an example where domain knowledge is incorporated into financial data analysis (the investment analysis), by means of portfolio cuts; (ii) London underground transportation system. The latter example demonstrates how graph theory can be used to identify the stations in the London underground network which have the greatest influence on the functionality of the traffic, and also to assess the impact of a station closure on service levels across the city.

References

Atwood, J. and D. Towsley (2016). "Diffusion-convolutional neural networks". In: *Advances in Neural Information Processing Systems*. 1993–2001.

Baba, K., R. Shibata, and M. Sibuya (2004). "Partial correlation and conditional correlation as measures of conditional independence". *Australian & New Zealand Journal of Statistics*. 46(4): 657–664.

Bacciu, D. and L. Di Sotto (2019). "A non-negative factorization approach to node pooling in graph convolutional neural networks". In: *Proceedings of the International Conference of the Italian Association for Artificial Intelligence*. Springer. 294–306.

Bacciu, D., F. Errica, and A. Micheli (2018). "Contextual graph Markov model: A deep and generative approach to graph processing". *arXiv preprint arXiv:1805.10636*.

Bacciu, D. and D. P. Mandic (2020). "Tensor decompositions in deep learning". In: *Proc. of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN'20)*. 441–450.

Baingana, B. and G. B. Giannakis (2016). "Tracking switched dynamic network topologies from information cascades". *IEEE Transactions on Signal Processing*. 65(4): 985–997.

Banerjee, O., L. E. Ghaoui, and A. d'Aspremont (2008). "Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data". *Journal of Machine Learning Research*. 9(Mar): 485–516.

Belkin, M. and P. Niyogi (2003). "Laplacian eigenmaps for dimensionality reduction and data representation". *Neural Computation*. 15(6): 1373–1396.

Belkin, M. and P. Niyogi (2008). "Towards a theoretical foundation for Laplacian-based manifold methods". *Journal of Computer and System Sciences*. 74(8): 1289–1308.

Berge, C. (1984). *Hypergraphs: Combinatorics of Finite Sets*. Vol. 45. Elsevier.

Black, F. and R. Litterman (1992). "Global portfolio optimization". *Financial Analysts Journal*. 48(5): 280–291.

Boginski, V., S. Butenko, and P. M. Pardalos (2003). "On structural properties of the market graph". In: *Innovations in Financial and Economic Networks*. Ed. by A. Nagurney. Edward Elgar Publishers. 29–45.

Boginski, V., S. Butenko, and P. M. Pardalos (2005). "Statistical analysis of financial networks". *Computational Statistics & Data Analysis*. 48(2): 431–443.

Boginski, V., S. Butenko, and P. M. Pardalos (2006). "Mining market data: A network approach". *Computers & Operations Research*. 33(11): 3171–3184.

Boginski, V., S. Butenko, S. O., S. Trunkhanov, and J. Gil Lafuente (2014). "A network-based data mining approach to portfolio selection via weighted clique relaxations". *Annals of Operations Research*. 216: 23–34.

Bohannon, A. W., B. M. Sadler, and R. V. Balan (2019). "A filtering framework for time-varying graph signals". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 341–376.

Bojchevski, A., O. Shchur, D. Zügner, and S. Günnemann (2018). "NetGAN: Generating graphs via random walks". *arXiv preprint arXiv:1803.00816*.

Brandes, U. (2005). *Network Analysis: Methodological Foundations*. Springer.

Bruna, J., W. Zaremba, A. Szlam, and Y. LeCun (2013). "Spectral networks and locally connected networks on graphs". *arXiv preprint arXiv:1312.6203*.

Caetano, T. S., J. J. McAuley, L. Cheng, Q. V. Le, and A. J. Smola (2009). "Learning graph matching". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 31(6): 1048–1058.

Calkin, N. J. and M. Lopez de Prado (2014a). "Stochastic flow diagrams". *Algorithmic Finance*. 3(1–2): 21–42.

Calkin, N. J. and M. Lopez de Prado (2014b). "The topology of macro financial flows: An application of stochastic flow diagrams". *Algorithmic Finance*. 3(1): 43–85.

Calkin, N. J. and M. Lopez de Prado (2016). "Building diversified portfolios that outperform out of sample". *The Journal of Portfolio Management*. 42(4): 59–69.

Calvi, G. G., A. Moniri, M. Mahfouz, Q. Zhao, and D. P. Mandic (2019). "Compression and interpretability of deep neural networks via Tucker tensor layer: From first principles to tensor valued back-propagation". *arXiv preprint arXiv:1903.06133*.

Camponogara, E. and L. F. Nazari (2015). "Models and algorithms for optimal piecewise-linear function approximation". *Mathematical Problems in Engineering*. 2015.

Candès, E. J., J. Romberg, and T. Tao (2006). "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". *IEEE Transactions on Information Theory*. 52(2): 489–509.

Chami, I., S. Abu-El-Haija, B. Perozzi, C. Re, and K. Murphy (2020). "Machine learning on graphs: A model and comprehensive taxonomy". *arXiv preprint arXiv:2005.03675*.

Chen, G., D. R. Glen, Z. S. Saad, J. P. Hamilton, M. E. Thomason, I. H. Gotlib, and R. W. Cox (2011). "Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis". *Computers in Biology and Medicine*. 41(12): 1142–1155.

Chen, S., A. Sandryhaila, J. M. Moura, and J. Kovačević (2015). "Signal recovery on graphs: Variation minimization". *IEEE Transactions on Signal Processing*. 63(17): 4609–4624.

Chen, S., R. Varma, A. Singh, and J. Kovačević (2016). "Signal recovery on graphs: Fundamental limits of sampling strategies". *IEEE Transactions on Signal and Information Processing Over Networks*. 2(4): 539–554.

Chepuri, S. P., S. Liu, G. Leus, and A. O. Hero (2017). "Learning sparse graphs under smoothness prior". In: *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 6508–6512.

Cichocki, A., N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic (2016). "Tensor networks for dimensionality reduction and large-scale optimization. Part 1: Low-rank tensor decompositions". *Foundations and Trends® in Machine Learning*. 9(4–5): 249–429.

Cichocki, A., A.-H. Phan, Q. Zhao, N. Lee, I. V. Oseledets, M. Sugiyama, and D. Mandic (2017). "Tensor networks for dimensionality reduction and large-scale optimization. Part 2: Applications and future perspectives". *Foundations and Trends® in Machine Learning*. 9(6): 431–673.

Cioacă, T., B. Dumitrescu, and M.-S. Stupariu (2019). "Graph-based wavelet multiresolution modeling of multivariate terrain data". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 479–507.

Clarke, R., H. De Silva, and S. Thorley (2002). "Portfolio constraints and the fundamental law of active management". *Financial Analysts Journal*. 58: 48–66.

Cooper, J. and A. Dutle (2012). "Spectra of uniform hypergraphs". *Linear Algebra and Its Applications*. 436(9): 3268–3292.

Dai, H., Z. Kozareva, B. Dai, A. Smola, and L. Song (2018). "Learning steady-states of iterative algorithms over graphs". In: *Proceedings of the International Conference on Machine Learning*. 1114–1122.

Dal Col, A., P. Valdivia, F. Petronetto, F. Dias, C. T. Silva, and L. G. Nonato (2019). "Wavelet-based visual data exploration". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 459–478.

Das, A., A. L. Sampson, C. Lainscsek, L. Muller, W. Lin, J. C. Doyle, S. S. Cash, E. Halgren, and T. J. Sejnowski (2017). "Interpretation of the precision matrix and its application in estimating sparse brain connectivity during sleep spindles from human electrocorticography recordings". *Neural Computation*. 29(3): 603–642.

De Cao, N. and T. Kipf (2018). "MolGAN: An implicit generative model for small molecular graphs". *arXiv preprint arXiv:1805.11973*.

De Miguel, V., L. Garlappi, and R. R. Uppal (2009). "Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy?" *Review of Financial Studies*. 22: 1915–1953.

Dees, B. S., A. G. Constantinides, and D. P. Mandic (2019). "Graph theory and metro traffic modelling". *arXiv preprint arXiv:1912.05964*.

Defferrard, M., X. Bresson, and P. Vandergheynst (2016). "Convolutional neural networks on graphs with fast localized spectral filtering". In: *Advances in Neural Information Processing Systems*. 3844–3852.

Dempster, A. P. (1972). "Covariance selection". *Biometrics*: 157–175.

Doersch, C. (2016). "Tutorial on variational autoencoders". *arXiv preprint arXiv:1606.05908*.

Dong, X., D. Thanou, P. Frossard, and P. Vandergheynst (2015). "Learning graphs from signal observations under smoothness prior". June [online]. Available: URL: http://arXiv.org/abs/1406.7842.

Dong, X., D. Thanou, P. Frossard, and P. Vandergheynst (2016). "Learning Laplacian matrix in smooth graph signal representations". *IEEE Transactions on Signal Processing*. 64(23): 6160–6173.

Dong, X., D. Thanou, M. Rabbat, and P. Frossard (2019). "Learning graphs from data: A signal representation perspective". *IEEE Signal Processing Magazine*. 36(3): 44–63.

Dorfler, F. and F. Bullo (2012). "Kron reduction of graphs with applications to electrical networks". *IEEE Transactions on Circuits and Systems I: Regular Papers*. 60(1): 150–163.

Epskamp, S. and E. I. Fried (2018). "A tutorial on regularized partial correlation networks". *Psychological Methods*. 23(4): 617.

Fiedler, M. (1973). "Algebraic connectivity of graphs". *Czechoslovak Mathematical Journal*. 23(2): 298–305.

Freeman, L. C. (1977). “A set of measures of centrality based on betweenness”. *Sociometry*. 40: 35–41.

Friedman, J., T. Hastie, and R. Tibshirani (2008). “Sparse inverse covariance estimation with the graphical LASSO”. *Biostatistics*. 9(3): 432–441.

Gauvin, L., A. Panisson, and C. Cattuto (2014). “Detecting the community structure and activity patterns of temporal networks: A non-negative tensor factorization approach”. *PLOS ONE*. 9(1): e86028.

Giannakis, G. B., Y. Shen, and G. V. Karanikolas (2018). “Topology identification and learning over graphs: Accounting for nonlinearities and dynamics”. *Proceedings of the IEEE*. 106(5): 787–807.

Gilmer, J., S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl (2017). “Neural message passing for quantum chemistry”. In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. 1263–1272.

Gori, M., G. Monfardini, and F. Scarselli (2005). “A new model for learning in graph domains”. In: *Proceedings of the IEEE International Joint Conference on Neural Networks, 2005*. Vol. 2. 729–734.

Grotas, S., Y. Yakoby, I. Gera, and T. Routtenberg (2019). “Power systems topology and state estimation by graph blind source separation”. *IEEE Transactions on Signal Processing*. 67(8): 2036–2051.

Grover, A. and J. Leskovec (2016). “Node2Vec: Scalable feature learning for networks”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 855–864.

Grover, A., A. Zweig, and S. Ermon (2019). “Graphite: Iterative generative modeling of graphs”. In: *Proc. International Conference on Machine Learning*. 2434–2444.

Gu, Y. and X. Wang (2019). “Local-set-based graph signal sampling and reconstruction”. In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 255–292.

Gunawardena, A. A., R. R. Meyer, and W. L. Dougan (2012). “Optimal selection of an independent set of cliques in a market graph”. In: *Proceedings of the International Conference on Economics, Business and Marketing Management*. 281–285.

Hagen, L. and A. B. Kahng (1992). "New spectral methods for ratio cut partitioning and clustering". *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 11(9): 1074–1085.

Hamilton, W., Z. Ying, and J. Leskovec (2017). "Inductive representation learning on large graphs". In: *Advances in Neural Information Processing Systems*. 1024–1034.

Hammond, D. K., P. Vandergheynst, and R. Gribonval (2011). "Wavelets on graphs via spectral graph theory". *Applied and Computational Harmonic Analysis*. 30(2): 129–150.

Hamon, R., P. Borgnat, P. Flandrin, and C. Robardet (2019). "Transformation from graphs to signals and back". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 111–139.

Hasanzadeh, A., E. Hajiramezanali, K. Narayanan, N. Duffield, M. Zhou, and X. Qian (2019). "Semi-implicit graph variational autoencoders". In: *Advances in Neural Information Processing Systems*. 10712–10723.

Ioannidis, V. N., D. Berberidis, and G. B. Giannakis (2019a). "GraphSAC: Detecting anomalies in large-scale graphs". *arXiv preprint arXiv:1910.09589*.

Ioannidis, V. N., Y. Shen, and G. B. Giannakis (2019b). "Semi-blind inference of topologies and dynamical processes over dynamic graphs". *IEEE Transactions on Signal Processing*. 67(9): 2263–2274.

Jeh, G. and J. Widom (2002). "SIMRANK: A measure of structural-context similarity". In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 538–543.

Kalofolias, V. (2016). "How to learn a graph from smooth signals". In: *Proceedings of the Artificial Intelligence and Statistics*. 920–929.

Kalyagin, V., A. Koldanov, P. Koldanov, and V. Zamaraev (2014). "Market graph and Markowitz model". In: *Optimization in Science and Engineering*. Ed. by T. M. Rassias, C. A. Floudas, and S. Butenko. Springer. 293–306.

Kaplan, D. (2008). *Structural Equation Modeling: Foundations and Extensions*. Vol. 10. Sage Publications.

Katsimpras, G. and G. Paliouras (2020). "Class-aware tensor factorization for multi-relational classification". *Information Processing & Management*. 57(2): 102068.

Kingma, D. P. and M. Welling (2013). "Auto-encoding variational Bayes". *arXiv preprint arXiv:1312.6114*.

Kipf, T. N. and M. Welling (2016a). "Semi-supervised classification with graph convolutional networks". *arXiv preprint arXiv:1609.02907*.

Kipf, T. N. and M. Welling (2016b). "Variational graph auto-encoders". *arXiv preprint arXiv:1611.07308*.

Kolaczyk, E. D. (2009). *Statistical Analysis of Network Data – Methods and Models*. New York: Springer-Verlag.

Kolm, P. N., R. Tutuncu, and F. J. Fabozzi (2014). "60 years of portfolio optimization: Practical challenges and current trends". *European Journal of Operational Research*. 234(2): 356–371.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner (1998). "Gradient-based learning applied to document recognition". *Proceedings of the IEEE*. 86(11): 2278–2324.

Ledoit, O. and M. Wolf (2003). "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection". *Journal of Empirical Finance*. 10(5): 603–621.

Li, S., Z. Yu, M. Xiang, and D. Mandic (2020). "Reciprocal adversarial learning via characteristic functions". *arXiv preprint arXiv:2006.08413*.

Li, Y., X. F. Jiang, Y. Tian, S. P. Li, and B. Zheng (2019). "Portfolio optimization based on network topology". *Physica A*. 515: 671–681.

Li, Y., D. Tarlow, M. Brockschmidt, and R. Zemel (2015). "Gated graph sequence neural networks". *arXiv preprint arXiv:1511.05493*.

Li, Y., O. Vinyals, C. Dyer, R. Pascanu, and P. Battaglia (2018). "Learning deep generative models of graphs". *arXiv preprint arXiv:1803.03324*.

Liben-Nowell, D. and J. Kleinberg (2007). "The link-prediction problem for social networks". *Journal of the American Society for Information Science and Technology*. 58(7): 1019–1031.

Lin, Y. R., Y. Chi, S. Zhu, H. Sundaram, and B. L. Tseng (2008). "Facetnet: A framework for analyzing communities and their evolutions in dynamic networks". In: *Proceedings of the International Conference on World Wide Web (WWW)*. 685–694.

Lin, Y., J. Sun, P. Castro, R. Konuru, H. Sundaram, and A. Kelliher (2009). "MetaFac: Community discovery via relational hypergraph factorization". In: *Proceedings of the ACM KDD International Conference on Knowledge Discovery and Data Mining*. 527–536.

Ma, T., J. Chen, and C. Xiao (2018). "Constrained generation of semantically valid graphs via regularizing variational autoencoders". In: *Advances in Neural Information Processing Systems*. 7113–7124.

Mandic, D. (2007). "Machine learning and signal processing applications of fixed point theory". *Tutorial in IEEE ICASSP, 2007*.

Mandic, D. and J. Chambers (2001). *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*. Wiley.

Mandic, D. P. and V. S. L. Goh (2009). *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. John Wiley & Sons.

Mao, X. and Y. Gu (2019). "Time-varying graph signals reconstruction". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 293–316.

Markowitz, H. (1952). "Portfolio selection". *Journal of Finance*. 7(1): 77–91.

Marques, A., A. Ribeiro, and S. Segarra (2017). "Graph signal processing: Fundamentals and applications to diffusion processes". In: *Proc. of the IEEE Int. Conf. Accoustic, Speech and Signal Processing, (ICASSP), Tutorial, 2017*.

Masuda, N., M. A. Porter, and R. Lambiotte (2017). "Random walks and diffusion on networks". *Physics Reports*. 716: 1–58.

Mateos, G., S. Segarra, A. G. Marques, and A. Ribeiro (2019). "Connecting the dots: Identifying network structure via graph signal processing". *IEEE Signal Processing Magazine*. 36(3): 16–43.

Mei, J. and J. M. Moura (2016). "Signal processing on graphs: Causal modeling of unstructured data". *IEEE Transactions on Signal Processing*. 65(8): 2077–2092.

Meinshausen, N. and P. Bühlmann (2006). "High-dimensional graphs and variable selection with the LASSO". *The Annals of Statistics*. 34(3): 1436–1462.

Micheli, A. (2009). "Neural network for graphs: A contextual constructive approach". *IEEE Transactions on Neural Networks*. 20(3): 498–511.

Monti, F., D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein (2017). "Geometric deep learning on graphs and manifolds using mixture model CNNs". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5115–5124.

Motl, J. and O. Schulte (2015). "The CTU prague relational learning repository". *arXiv preprint arXiv:1511.03086*.

Ng, A. Y., M. I. Jordan, and Y. Weiss (2002). "On spectral clustering: Analysis and an algorithm". *Advances in Neural Information Processing Systems*. 2002: 849–856.

Nickel, M., V. Tresp, and H.-P. Kriegel (2011). "A three-way model for collective learning on multi-relational data". In: *Proceedings of the 28th International Conference on Machine Learning*. 809–816.

Novikov, A., D. Podoprikhin, A. Osokin, and D. P. Vetrov (2015). "Tensorizing neural networks". In: *Advances in Neural Information Processing Systems (NIPS)*. 442–450.

Pan, S., R. Hu, G. Long, J. Jiang, L. Yao, and C. Zhang (2018). "Adversarially regularized graph autoencoder for graph embedding". *arXiv preprint arXiv:1802.04407*.

Papalexakis, E. E., L. Akoglu, and D. Lence (2013). "Do more views of a graph help? Community detection and clustering in multi-graphs". In: *Proceedings of the 16th International Conference on Information Fusion*. 899–905.

Pasdeloup, B., V. Gripon, R. Alami, and M. G. Rabbat (2019). "Uncertainty principle on graphs". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 317–340.

Pavez, E. and A. Ortega (2016). "Generalized Laplacian precision matrix estimation for graph signal processing". In: *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016*. 6350–6354.

Peralta, G. and A. Zareei (2016). “A network approach to portfolio selection”. *Journal of Empirical Finance*. 38(A): 157–180.

Perozzi, B., R. Al-Rfou, and S. Skiena (2014). “Deepwalk: Online learning of social representations”. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 701–710.

Pourahmadi, M. (2011). “Covariance estimation: The GLM and regularization perspectives”. *Statistical Science*: 369–387.

Rabiei, H., F. Richard, O. Coulon, and J. Lefèvre (2019). “Estimating the complexity of the cerebral cortex folding with a local shape spectral analysis”. In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 437–458.

Raffinot, T. (2017). “Hierarchical clustering-based asset allocation”. *The Journal of Portfolio Management*. 44(2): 89–99.

Sadhanala, V., Y.-X. Wang, and R. Tibshirani (2016). “Graph sparsification approaches for Laplacian smoothing”. In: *Artificial Intelligence and Statistics*. 1250–1259.

Saito, S., D. P. Mandic, and H. Suzuki (2018). “Hypergraph p-Laplacian: A differential geometry view”. In: *Proc. of the Thirty-Second AAAI Conference on Artificial Intelligence*. 3984–3991.

Sakiyama, A., Y. Tanaka, T. Tanaka, and A. Ortega (2019). “Eigendecomposition-free sampling set selection for graph signals”. *IEEE Transactions on Signal Processing*. 67(10): 2679–2692.

Scalzo, B., L. Stanković, A. G. Constantinides, and D. P. Mandic (2020). “Portfolio cuts: A graph-theoretic framework to diversification”. In: *Proc. of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 8454–8458.

Scarselli, F., M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini (2008). “The graph neural network model”. *IEEE Transactions on Neural Networks*. 20(1): 61–80.

Schaeffer, S. E. (2007). “Graph clustering”. *Computer Science Review*. 1(1): 27–64.

Segarra, S., A. G. Marques, G. Mateos, and A. Ribeiro (2016). “Blind identification of graph filters with multiple sparse inputs.” In: *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 4099–4103.

Segarra, S., A. G. Marques, G. Mateos, and A. Ribeiro (2017). "Network topology inference from spectral templates". *IEEE Transactions on Signal and Information Processing Over Networks*. 3(3): 467–483.

Sen, P., G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad (2008). "Collective classification in network data". *AI Magazine*. 29(3): 93–93.

Shi, J. and J. Malik (2000). "Normalized cuts and image segmentation". *Departmental Papers (CIS)*: 107.

Simon, H. A. (1991). "The architecture of complexity". In: *Facets of Systems Science*. Springer. 457–476.

Slawski, M. and M. Hein (2015). "Estimation of positive definite M-matrices and structure learning for attractive Gaussian Markov random fields". *Linear Algebra and Its Applications*. 473: 145–179.

Spielman, D. A. and S. H. Teng (2007). "Spectral partitioning works: Planar graphs and finite element meshes". *Linear Algebra and its Applications*. 421(2–3): 284–305.

Stanković, L. (2001). "A measure of some time–frequency distributions concentration". *Signal Processing*. 81(3): 621–631.

Stanković, L. (2015). *Digital Signal Processing with Selected Topics*. CreateSpace Independent Publishing Platform, An Amazon.com Company.

Stanković, L., M. Daković, D. Mandic, M. Brajović, B. Scalzo, and A. Constantinides (2020). "A low-dimensionality method for data-driven graph learning". In: *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 5340–5344.

Stanković, L., M. Daković, and E. Sejdić (2017). "Vertex-frequency analysis: A way to localize graph spectral components [Lecture Notes]". *IEEE Signal Processing Magazine*. 34(4): 176–182.

Stanković, L., D. P. Mandic, M. Daković, M. Brajović, B. Scalzo, and T. Constantinides (2019a). "Graph signal processing – Part I: Graphs, graph spectra, and spectral clustering". *arXiv:1907.03467*.

Stanković, L., D. P. Mandic, M. Dakovic, I. Kisil, E. Sejdic, and A. G. Constantinides (2019b). "Understanding the basis of graph signal processing via an intuitive example-driven approach [Lecture Notes]". *IEEE Signal Processing Magazine*. 36(6): 133–145.

Stanković, L., D. P. Mandic, M. Dakovic, and I. Kisil (2020). "Demystifying the coherence index in compressive sensing [Lecture Notes]". *IEEE Signal Processing Magazine*. 37(1): 152–162.

Stanković, L. and E. Sejdić (2019). *Vertex-Frequency Analysis of Graph Signals*. Springer.

Stanković, L., E. Sejdić, and M. Daković (2018). "Reduced interference vertex-frequency distributions". *IEEE Signal Processing Letters*. 25(9): 1393–1397.

Stanković, L., E. Sejdić, S. Stanković, M. Daković, and I. Orović (2019c). "A tutorial on sparse signal reconstruction and its applications in signal processing". *Circuits, Systems, and Signal Processing*. 38(3): 1206–1263.

Stoer, M. and F. Wagner (1997). "A simple min-cut algorithm". *Journal of the ACM (JACM)*. 44(4): 585–591.

Tanaka, Y. and Y. C. Eldar (2019). "Generalized sampling on graphs with subspace and smoothness priors". *arXiv preprint arXiv:1905.04441*.

Tanaka, Y. and A. Sakiyama (2019). "Oversampled transforms for graph signals". In: *Vertex-Frequency Analysis of Graph Signals*. Ed. by L. Stanković and E. Sejdić. Springer. 223–254.

Tang, J., M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei (2015). "Line: Large-scale information network embedding". In: *Proceedings of the 24th International Conference on World Wide Web*. 1067–1077.

Tang, L. and H. Liu (2011). "Leveraging social media networks for classification". *Data Mining and Knowledge Discovery*. 23(3): 447–478.

Tang, W., Z. Lu, and I. S. Dhillon (2009). "Clustering with multiple graphs". In: *Proceedings of the Ninth IEEE International Conference on Data Mining*. 1016–1021.

Thanou, D., D. I. Shuman, and P. Frossard (2014). "Learning parametric dictionaries for signals on graphs". *IEEE Transactions Signal Processessing*. 62(15): 3849–3862.

Thanou, D., X. Dong, D. Kressner, and P. Frossard (2017). "Learning heat diffusion graphs". *IEEE Transactions on Signal and Information Processing Over Networks*. 3(3): 484–499.

Transport for London (n.d.). URL: https://tfl.gov.uk/.

Ubaru, S., J. Chen, and Y. Saad (2017). "Fast estimation of tr(f(A)) via stochastic Lanczos quadrature". *SIAM Journal on Matrix Analysis and Applications*. 38(4): 1075–1099.

Veličković, P., G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio (2017). "Graph attention networks". *arXiv preprint arXiv:1710.10903*.

Verma, A. and K. K. Bharadwaj (2017a). "A comparative study based on tensor factorization and clustering techniques for community mining in heterogeneous social network". *In Proceedings of the International Conference on Computing, Communication and Networking Technologies (ICCCNT)*: 1–6.

Verma, A. and K. K. Bharadwaj (2017b). "Identifying community structure in a multi-relational network employing non-negative tensor factorization and GA k-means clustering". *Wires: Data Mining and Knowledge Discovery*. 7(1): 1–32.

Von Luxburg, U. (2007). "A tutorial on spectral clustering". *Statistics and Computing*. 17(4): 395–416.

Wagner, D. and F. Wagner (1993). "Between min cut and graph bisection". In: *Proc. of the International Symposium on Mathematical Foundations of Computer Science*. Springer. 744–750.

Wai, H.-T., Y. C. Eldar, A. E. Ozdaglar, and A. Scaglione (2019). "Community inference from graph signals with hidden nodes". In: *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 4948–4952.

Wang, H., J. Wang, J. Wang, M. Zhao, W. Zhang, F. Zhang, X. Xie, and M. Guo (2017). "GraphGAN: Graph representation learning with generative adversarial nets". *arXiv preprint arXiv:1711.08267*.

Wang, Z., A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli (2004). "Image quality assessment: From error visibility to structural similarity". *IEEE Transactions on Image Processing*. 13(4): 600–612.

Wu, Z., S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu (2019). "A comprehensive survey on graph neural networks". *arXiv preprint arXiv:1901.00596*.

Xu, Y. L. and D. P. Mandic (2020). "Recurrent graph tensor networks". *arXiv preprint arXiv:2009.08727*. Sept. arXiv: 2009.08727 [cs.LG].

Xu, Y. L., K. Konstantinidis, and D. P. Mandic (2020). "Multi-graph tensor networks". In: *Advances in Neural Information Processing Systems*.

Yankelevsky, Y. and M. Elad (2016). "Dual graph regularized dictionary learning". *IEEE Transactions on Signal and Information Processing Over Networks*. 2(4): 611–624.

You, J., R. Ying, X. Ren, W. L. Hamilton, and J. Leskovec (2018). "GraphRNN: Generating realistic graphs with deep auto-regressive models". *arXiv preprint arXiv:1802.08773*.

Yu, W., C. Zheng, W. Cheng, C. C. Aggarwal, D. Song, B. Zong, H. Chen, and W. Wang (2018). "Learning deep network representations with adversarially regularized autoencoders". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2663–2671.

Yuan, M. and Y. Lin (2006). "Model selection and estimation in regression with grouped variables". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 68(1): 49–67.

Yuan, M. and Y. Lin (2007). "Model selection and estimation in the Gaussian graphical model". *Biometrika*. 94(1): 19–35.

Zhang, J., X. Shi, J. Xie, H. Ma, I. King, and D.-Y. Yeung (2018a). "GAAN: Gated attention networks for learning on large and spatiotemporal graphs". *arXiv preprint arXiv:1803.07294*.

Zhang, J., X. Shi, S. Zhao, and I. King (2019a). "STAR-GCN: Stacked and reconstructed graph convolutional networks for recommender systems". *arXiv preprint arXiv:1905.13129*.

Zhang, M. and Y. Chen (2018). "Link prediction based on graph neural networks". In: *Advances in Neural Information Processing Systems*. 5165–5175.

Zhang, M., S. Jiang, Z. Cui, R. Garnett, and Y. Chen (2019b). "D-VAE: A variational autoencoder for directed acyclic graphs". In: *Advances in Neural Information Processing Systems*. 1588–1600.

Zhang, Z., P. Cui, and W. Zhu (2018b). "Deep learning on graphs: A survey". *arXiv preprint arXiv:1812.04202*.

Zhao, T., H. Liu, K. Roeder, J. Lafferty, and L. Wasserman (2012). "The huge package for high-dimensional undirected graph estimation in R". *Journal of Machine Learning Research*. 13(Apr): 1059–1062.

Zheng, M., J. Bu, C. Chen, C. Wang, L. Zhang, G. Qiu, and D. Cai (2011). "Graph regularized sparse coding for image representation". *IEEE Transactions on Image Processing*. 20(5): 1327–1336.

Zhou, D., J. Huang, and B. Schölkopf (2007). "Learning with hypergraphs: Clustering, classification, and embedding". In: *Advances in Neural Information Processing Systems*. 1601–1608.

Zhou, J., G. Cui, Z. Zhang, C. Yang, Z. Liu, and M. Sun (2018). "Graph neural networks: A review of methods and applications". *arXiv preprint arXiv:1812.08434*.

Zhu, X. J. (2005). "Semi-supervised learning literature survey". *Tech. rep.* University of Wisconsin-Madison, Department of Computer Sciences.