Summary of "U-statistics of row-column exchangeable matrices: application to ecological network analysis"

Tâm Le Minh*

*Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France

Abstract: *This is a summary of my doctoral thesis [2], which I have defended on November 13th, 2023. I have been supervised by Stéphane Robin, Sophie Donnet and François Massol.*

The work presented in this thesis is essentially theoretical [3, 5, 4], but motivated by ecological applications [1]. Ecological interaction networks represent the functioning of an ecosystem. Investigating the variability of interaction networks enables us to understand how the ecosystems are affected by external factors. This thesis suggests a methodology to analyze bipartite networks, applicable to ecological mutualistic networks.

This methodology is based on U-statistics of row-column exchangeable matrices. Row-column exchangeable matrices are random matrices, the joint probability distribution of which is invariant by simultaneous permutations of rows and columns. U-statistics correspond to the class of statistics defined as the empirical mean of a function of a subset, over all subsets of observations. U-statistics of matrices are the average of a submatrix function over the entire matrices. In network analysis, row-column exchangeable matrices are the adjacency matrices of bipartite node-exchangeable networks, and U-statistics can be used as estimators of quantities of interest.

This thesis focuses on the asymptotic behavior of the U-statistics of row-column exchangeable matrices. In the first part, backward martingales are used to derive a limit theorem on U-statistics of row-column exchangeable matrices. In the second part, a Hoeffding-type decomposition is established for them, which extends the previous limit theorem. Inspired by this decomposition, an estimator of the asymptotic variance is also suggested, making it possible to propose a general method for performing statistical inference tasks on exchangeable network models. The third part of the thesis extends the methodology to degenerate U-statistics, which have a faster rate of convergence.

These statistical developments are applied to the analysis of bipartite networks, including mutualistic ecological networks. Many ecological questions are interested in the general structure of networks rather than the collection of present species. This makes exchangeable random network models, the adjacency matrices of which are row-column exchangeable, well-suited to analyze these networks. U-statistics are used as estimators of quantities of interest such as the degree heterogeneity, motif densities, or graphon metrics. It is possible to obtain statistical guarantees on these estimators, for example in the form of confidence intervals, owing to the theoretical results and the methodology developed in this thesis. Some examples of exchangeable random network models and U-statistics are given, answering real

ecological questions. Simulation studies are used to validate the use of this methodology for these examples.

References

[1] M. Dubart, P. Alonso, D. Barroso-Bergada, N. Becker, K. Bethune, D. A. Bohan, C. Boury, M. Cambon, E. Canard, E. Chancerel, J. Chiquet, P. David, N. de Manincor, S. Donnet, A. Duputié, B. Facon, E. Guichoux, T. Le Minh, S. Ortiz-Martínez, L. Piouceau, A. Sacco-Martret de Préville, M. Plantegenest, C. Poux, V. Ravigné, S. Robin, M. Trillat, C. Vacher, C. Vernière, and F. Massol. Coupling ecological network analysis with high-throughput sequencing-based surveys: Lessons from the next-generation biomonitoring project. In *The Future of Agricultural Landscapes, Part III*, volume 65 of *Advances in Ecological Research*, pages 367–430. Academic Press, 2021.

[2] T. Le Minh. *U-statistics of row-column exchangeable matrices: application to ecological network analysis*. PhD thesis, Université Paris-Saclay, 2023.

[3] T. Le Minh. U-statistics on bipartite exchangeable networks. *ESAIM: Probability and Statistics*, 27:576–620, 2023.

[4] T. Le Minh. Characterization of the asymptotic behavior of U-statistics on row-column exchangeable matrices. *arXiv preprint arXiv:2401.07876*, 2024.

[5] T. Le Minh, S. Donnet, F. Massol, and S. Robin. Hoeffding-type decomposition for U-statistics on bipartite networks. *arXiv preprint arXiv:2308.14518*, 2023.