# EXCHANGEABLE RANDOM GRAPHS

VICTOR VEITCH

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy Graduate Department of Statistical Sciences University of Toronto

May 2017 – version 4.2

As far as I can see, all a priori statements in physics have their origin in symmetry.

— H. Weyl

# ABSTRACT


Short summary of the contents in English... a great guide by Kent Beck how to write good abstracts can be found here:

`https://plg.uwaterloo.ca/~migod/research/beckOOPSLA.html`

# PUBLICATIONS

This might come in handy for PhD theses: some ideas and figures have appeared previously in the following publications:

*This is just an early – and currently ugly – test!*

*Attention*: This requires a separate run of `bibtex` for your `refsection`, e. g., `ClassicThesis1-blx` for this file. You might also use `biber` as the backend for `biblatex`. See also `http://tex.stackexchange.com/questions/128196/problem-with-refsection`.

*We have seen that computer programming is an art,*
*because it applies accumulated knowledge to the world,*
*because it requires skill and ingenuity, and especially*
*because it produces objects of beauty.*

**— knuth:1974 [knuth:1974]**

# ACKNOWLEDGMENTS

Put your acknowledgments here.

Many thanks to everybody who already sent me a postcard!

Regarding the typography and other help, many thanks go to Marco Kuhlmann, Philipp Lehman, Lothar Schlesier, Jim Young, Lorenzo Pantieri and Enrico Gregorio[1], Jörg Sommer, Joachim Köstler, Daniel Gottschlag, Denis Aydin, Paride Legovini, Steffen Prochnow, Nicolas Repp, Hinrich Harms, Roland Winkler, Jörg Weber, Henri Menke, Claus Lahiri, Clemens Niederberger, Stefano Bragaglia, Jörn Hees, and the whole LaTeX-community for support, ideas and some great software.

*Regarding LyX*: The LyX port was intially done by *Nicholas Mariette* in March 2009 and continued by *Ivo Pletikosić* in 2011. Thank you very much for your work and for the contributions to the original style.

1 Members of GuIT (Gruppo Italiano Utilizzatori di TeX e LaTeX)

# CONTENTS


I OVERVIEW 1

II THE SHOWCASE 3

1 INTRODUCTION 5

- 1.0.1 The Class of Random Graphs Arising from Exchangeable Random Measures 7
- 1.0.2 Sampling and Estimation for (Sparse) Exchangeable Graphs 8
- 1.0.3 Sampling Perspectives on (Sparse) Exchangeable Graphs 8

BIBLIOGRAPHY 11

# LIST OF FIGURES

# LIST OF TABLES

# LISTINGS

# ACRONYMS

# Part I

# OVERVIEW

# Part II

# THE SHOWCASE

You can put some informational part preamble text here. Illo principalmente su nos. Non message *occidental* anglo-romanic da. Debitas effortio simplificate sia se, auxiliar summarios da que, se avantiate publicationes via. Pan in terra summarios, capital interlingua se que. Al via multo esser specimen, campo responder que da. Le usate medical addresses pro, europa origine sanctificate nos se.

# 1 INTRODUCTION

The topic of this thesis is the mathematical foundations for statistical modeling of relational or network data. In this setting, the dataset $g$ is modeled as a graph, idealized as a sample drawn from some probability distribution over graphs, $P(g;\theta)$, whose parameters are in some unknown configuration. The goal is then to infer the configuration of the parameters on the basis of the observation, and thus understand the properties of the observed network. We substantially develop the associated statistical theory. Along the way, we uncover some interesting insights about the foundations of statistical inference generally; particularly, on the role of probabilistic symmetries, sampling, empirical measures, and the interplay between these subjects.

The first main challenge is to identify families of probability distributions over graphs that are suitable for modeling network data. A satisfactory family of distributions should, minimally, meet the following basic desiderata:

1. samples from the distributions should reflect the rich structure of real-world networks of interest
2. it should be possible to recover the parameters of the observed network on the basis of the observed data

Remarkably, there is no general approach to the statistical modeling of network data that satisfies these requirements. More precisely, there is already well developed and powerful theory—the *graphon* framework—for the analysis of densely connected networks (that is, networks that contain a constant fraction of all possible edges as the size becomes large), but there is no comparable theory for sparse (that is, not dense) networks. The vast majority of real-world networks are sparsely connected, so the absence of a satisfactory modeling framework is a considerable omission. This thesis addresses the gap, generalizing the graphon framework to the sparse graph setting.

The most popular approaches to statistical network modeling, e.g. [NS01; HRH02; ABFX08; MJG09; LOGR12], fall under the remit of the graphon framework for dense random graphs. Graphon models are the class of generative models for random graphs where each vertex $i$ is assigned some independent latent features $\vartheta_i$, and, conditional on these latent features, each pair of vertices $i,j$ is connected by an edge independently with probability $W(\vartheta_i,\vartheta_j)$ determined by the latent features of $i$ and $j$; the function $W$ is known as a graphon. This is a very natural class of models, and models of this type, such as stochastic block models and latent feature models, have a long history in the literature. However, as already alluded to, graphon models give rise to random graphs that are almost surely *dense* (or trivially empty if the graphon is $W = 0$ almost everywhere). Thus, as stated plainly in [OR15], graphon models are misspecified as statistical models for most real-world data.

A key to arriving at the generalized models of this thesis is a second perspective on graphon models rooted in probabilistic symmetry. In this view, we identify a projective family $(G_n)_{n \in \mathbb{N}}$ of random graphs with the upper left $n \times n$ submatrices of an infinite random adjacency matrix $A$, and then define the class of models to be those such that the distribution of $A$ is invariant under joint permutations of its rows and columns. This *exchangeability* of $A$ is a natural formalization of the requirement that the labels of the vertices of a random graph should be uninformative about the structure of the graph. The fact that the graphon models are the models defined by exchangeability of the infinite adjacency matrix is, essentially, the content of the celebrated Aldous–Hoover theorem [Ald81; Hoo79].

This derivation of the dense graph framework is a particular instance of a general recipe for constructing statistical models: a probabilistic symmetry is assumed on some infinite random structure and an associated representation theorem characterizes the ergodic measures, forming the foundation of a framework for statistical analysis. We follow this broad strategy.

Our inspiration comes from recent paper of Caron and Fox [CF14] that exploits a bijection between random graphs and point processes on $\mathbb{R}_+^2$ to exhibit a class of sparse random graphs. In their paper, they observe that their random graphs satisfy a natural notion of exchangeability exchangeability when considered as a point process, and make use of an associated representation theorem to study the model. We reverse this chain of reasoning, beginning with the symmetry on point processes and elucidating the full family of random graphs that arise from the associated representation theorem.

Following Caron and Fox, we represent random graphs as an infinite simple point processes on $\mathbb{R}_+^2$ with finite random graphs given by truncating the support of the point process to a finite set (see **??**). Accordingly, each edge is identified with a label $(\theta_i, \theta_j) \in \mathbb{R}_+^2$, and each vertex with a label $\theta_i \in \mathbb{R}_+$. The representation theorem associated to exchangeable point processes is known by the work of Kallenberg [Kal90; Kal05]. We arrive at our representation theorem by a straightforward translation of this result into the random graph setting. The distribution of every random graph satisfying the symmetry is characterized by a triple $\mathcal{W} = (I, S, W)$ where $I \in \mathbb{R}_+$, $S : \mathbb{R}_+ \to \mathbb{R}_+$ is a measurable function such that $\min(S, 1)$ is integrable, and $W : \mathbb{R}_+^2 \to [0, 1]$ is a symmetric measurable function satisfying certain weak integrability conditions. (See **??**; $W$ integrable is sufficient but not necessary.) The associated random graphs have three possible components: isolated edges, governed by $I$, infinite stars, governed by $S$, and a final piece that provides the interesting graph structure, governed by $\mathcal{W}$.

We give the probability distribution $P(\Gamma; \mathcal{W})$ over (infinite) random graphs $\Gamma$ associated with each graphex in terms of the associated generative model. Let $\theta = \mathbb{R}_+$ be the space of labels of the graph, $\vartheta = \mathbb{R}_+$ be the space of latent parameters, and $\Pi = \{\theta_i, \vartheta_i\}_{i \in \mathbb{N}}$ be a unit rate Poisson process on $\theta \times \vartheta$. For all $i < j$, conditioned on the points of $\Pi$, the random graph includes the edge $\theta_i \leftrightarrow \theta_j$ with probability $W(\vartheta_i, \vartheta_j)$, independently of all other edges. This is the interesting component of the graph structure. The infinite stars component is given by, for each point $(\theta_i, \vartheta_i)$, including edges $\{(\theta_i, \sigma_{ij})\}$, where $\{\sigma_{ij}\}_{j \in \mathbb{N}}$ is distributed as a rate $S(\vartheta_i)$ Poisson process independent of everything else. The isolated edges component is given by including edges $\{(\rho_k, \rho'_k)\} \subset \theta^2$ where $\{(\rho_k, \rho'_k)\}$ are distributed as a rate $I$ Poisson process independent of everything else. To pass from an infinite random graph $\Gamma$ with vertices labeled in $\mathbb{R}_+$ to a projective sequence $(\Gamma_s)_{s \in \mathbb{R}_+}$ we define $\Gamma_s$ to be the (labeled) graph given by restricting to vertices with labels $\theta_i < s$, and removing any vertices that are isolated in the induced subgraph.

The contribution of this thesis may now be stated plainly: we introduce and develop the graphex framework for the statistical modeling of network and relational data.

The main content of the thesis is organized into three chapters on different aspects of the theoretical development. The three chapters correspond to the papers [VR15], [VR16] and [**Borgs:Chayes:Cohn:Veitch:2017**] respectively. In each case, the content of the chapters very closely follows the associated papers—the exceptions are minimal edits to ensure consistent notation and ease the flow of reading (by, e.g., not recapitulating the representation theorem three times).

### *The Class of Random Graphs Arising from Exchangeable Random Measures*

In **??**, we use the representation theorem to derive the graphex framework and we argue that the models of the framework are good candidates for the statistical modeling of real-world data. We first explain in detail the core idea of defining model classes by probabilistic symmetries, in particular, giving due attention to the key role of projectivity and the distinction between finite and infinite exchangeability. We then apply Kallenberg's representation theorem to derive the representation theorem for sparse exchangeable graphs, thereby defining the graphex framework. It remains to show that the models of graphex framework can capture the complex structure of real-world networks. Our first contribution to that end is to show that the dense graphon models are (essentially) contained as a special case of the graphex models. Next, we derive some sampling distribution properties of $(\Gamma_s)_{s \in \mathbb{R}_+}$ in terms of the generating graphex. This allows us to show that graphex models can indeed give rise to sparse graphs and, moreover, that they encompass much of the rich structure required for high fidelity modeling; e.g. they naturally encompass models with extreme degree heterogeneity ("power law" degree distributions), and very short typical path lengths even in the sparse regime ("small world" behaviour).

*Sampling and Estimation for (Sparse) Exchangeable Graphs*

In **??**, we identify the sampling scheme that is naturally associated with the models of the framework, and we introduce a general consistent estimator for the graphex. The sampling scheme is a modification of independent vertex sampling that throws away vertices that are isolated in the sampled subgraph. The estimator is the analogue of the empirical measure of classical i.i.d. statistics. The estimator is the dilation of the empirical graphon estimator, which is already known to be a consistent estimator in the dense graphon setting.

The results of this chapter are important for a number of reasons: Identifying the sampling scheme associated to the graphex models helps clarify the meaning and applicability of the models. Finding the analogue of the empirical measure in this setting opens the door to the development of non-parametric techniques for graphs. Less obviously, the machinery developed in order to formalize the estimation problem is critical for any further statistical theory; for example, pinning down the right notions of consistent estimation and identifiability in this setting is a subtle problem that is now resolved.

It also worth emphasizing that sampling and estimation and extremely tightly connected. One of the key ideas in the development of our estimator is that a sample from a very large random graph $\Gamma_s$ generated by $\mathcal{W}$ looks approximately like a (smaller) random graph drawn from $\mathcal{W}$ directly, even conditional on $\Gamma_s$. This is the network data analogue of the familiar idea that one can approximate samples from probability measure simply by subsampling a very large dataset. The validity of this approach, in both the sequence and network settings, relies critically on the exchangeability assumption—this is one way in which the tight connection between sampling and exchangeability manifests. Making this precise is one of the major contributions of this chapter.

*Sampling Perspectives on (Sparse) Exchangeable Graphs*

**??** addresses two closely related questions: When is it appropriate to model a network dataset using the graphex framework? And, what is the relationship in this setting between statistical modeling and limits of large sparse graph sequences?

The context for the later question is that, in parallel to [VR15], a group of graph theorists independently developed the graphex framework in the context of graph limit theories [BCCH16]. In the graph limit setting, the central problem is: what is a sensible notion of convergence for a sequence $G_1, G_2, \ldots$ of (non-random) graphs of growing size, and what is the natural limit object? In the dense graph setting, many superficially distinct notions of convergence turn out to coincide, with graphons as the natural limit object. See [**Lovasz:2012**] for a textbook level review. Indeed, the dense graphon models (and the term graphon) were developed by graph theorists independently of the statistical work on exchangeable random graph models. The two perspectives were unified in [DJ08; Aus08]; part of the contribution of **??** is to repeat this unification in the sparse graph setting.

Very informally, we define a new notion of graph limit—*sampling convergence*—by saying that a graph sequence $G_1, G_2, \ldots$ converges if randomly sampled subgraphs converge in distribution. More precisely, for each $G_k$ we sample $H_k^s$ by picking $\frac{s}{\sqrt{2e(G_k)}}$ vertices at random, and returning the induced subgraph with the vertices removed. The reason for picking each vertex with a probability proportional to $1/\sqrt{2e(G_k)}$ is that this keeps the expected number of edges of $H_k^s$ constant. We say the sequence $G_1, G_2, \ldots$ is sampling convergent if $H_k^s$ converges in distribution for all $s \in \mathbb{R}_+$. The main result is that whenever a sequence converges in this sense, the (weak) limit of $\Pr(H_k^s \in \cdot)$ is equal to the distribution of a size-s sample from some graphex $\mathcal{W}$. This gives another close connection between sampling and exchangeability (in fact, our proof of this result makes explicit use of Kallenberg's representation theorem). The connection to [BCCH16] is given by our demonstration that convergence in the sense of [BCCH16] implies sampling convergence.

In fact, we further show that every integrable graphex $\mathcal{W}$ arises as the limit of some sampling convergent graph sequence. This later observation has the nice consequence that we can use the limit theory to define the graphex models without any explicit appeal to the rather strange representation of graphs as point processes; it thus sidesteps questions about the interpretation of this representation and the associated notion of exchangeability. Instead, we can use the limit theory to formalize the idea that a graphex distribution is appropriate for modelling a network that can be thought of as generated by independently choosing nodes from a much larger network, and then throwing away isolated nodes. This idea was also at least implicit in the construction of the estimator of **??**.

This last observation begs the question of whether we could have defined the graphex models in terms of the sampling scheme in the first place, without any appeal to either exchangeability or graph limits. Indeed, this is the case. In **??** we show that graphex models have the property that if a graph $H_{s,p}$ is generated by first generating a size-

$s$ graph $\Gamma_s$ according to $\mathcal{W}$ and then independently including each vertex of $\Gamma_s$ with probability $p$, and finally throwing away the isolated vertices in the sample, then (ignoring labels) $H_{s,p}$ is distributed as a size $ps$ graph generated from $\mathcal{W}$. In **??** we show that this is a defining property of the graphex models. In particular, this sampling invariance is exactly equivalent to the exchangeability we use to define the graphex models.

[ABFX08] E. M. Airoldi, D. M. Blei, S. E. Feinberg, and E. P. Xing. "Mixed-membership stochastic blockmodels". In: *J. Machine Learn. Res.* 9 (2008), pp. 1981–2014.

[AG89] R. Arratia and L. Gordon. "Tutorial on large deviations for the binomial distribution". English. In: *Bulletin of Mathematical Biology* 51.1 (1989), pp. 125–131. ISSN: 0092-8240.

[Ald09] D. J. Aldous. "More Uses of Exchangeability: Representations of Complex Random Structures". In: *ArXiv e-prints* (Sept. 2009). arXiv: 0909.4339 [math.PR].

[Ald81] David J. Aldous. "Representations for partially exchangeable arrays of random variables". In: *J. Multivariate Anal.* 11.4 (1981), pp. 581–598. ISSN: 0047-259X.

[Aus08] Tim Austin. "On exchangeable random variables and the statistics of large graphs and hypergraphs". In: *Probab. Surv* 5.80-145 (2008).

[BA99] Albert-László Barabási and Réka Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (1999), pp. 509–512. eprint: http://www.sciencemag.org/content/286/5439/509.full.pdf.

[BBCS14] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. "Asymptotic behavior and distributional limits of preferential attachment graphs". In: *ArXiv e-prints* (Jan. 2014). arXiv: 1401.2792 [math.PR].

[BCCG15] C. Borgs, J. T. Chayes, H. Cohn, and S. Ganguly. "Consistent nonparametric estimation for heavy-tailed sparse graphs". In: *ArXiv e-prints* (Aug. 2015). arXiv: 1508.06675 [math.ST].

[BCCH16] C. Borgs, J. T. Chayes, H. Cohn, and N. Holden. "Sparse exchangeable graphs and their limits via graphon processes". In: *ArXiv e-prints* (Jan. 2016). arXiv: 1601.07134 [math.PR].

[BCCZ14a] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. "An $L^p$ theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions". In: *ArXiv e-prints* (Jan. 2014). arXiv: 1401.2906 [math.CO].

[BCCZ14b] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao. “An $L^p$ theory of sparse graph convergence II: LD convergence, quotients, and right convergence”. In: *ArXiv e-prints* (Aug. 2014). arXiv: 1408.0744 [math.CO].

[BCLS+06] C. Borgs, J. Chayes, L. Lovászsz, V. T. Sós, and K. Vesztergombi. “Counting graph homomorphisms”. In: *Topics in Discrete Mathematics*. Ed. by M. Klazar, J. Kratochvil, M. Loebl, J. Matousek, R. Thomas, and P. Valtr. Springer, 2006, pp. 315–371.

[BCLS+08a] C. Borgs, J. T. Chayes, L. Lovasz, V. T. Sós, and K. Vesztergombi. “Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing”. In: *Advances in Mathematics* 219.6 (2008), pp. 1801 –1851. ISSN: 0001-8708.

[BCLS+08b] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. “Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing”. In: *Adv. Math.* 219.6 (2008), pp. 1801–1851. ISSN: 0001-8708.

[BCLS+12] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. “Convergent sequences of dense graphs. II. Multiway cuts and statistical physics”. In: *Ann. Math.* 176 (2012), pp. 151–219.

[BCS15] C. Borgs, J. T. Chayes, and A. Smith. “Private Graphon Estimation for Sparse Graphs”. In: *ArXiv e-prints* (June 2015). arXiv: 1506.06162 [math.ST].

[BS01] Itai Benjamini and Oded Schramm. “Recurrence of distributional limits of finite planar graphs”. In: *Electron. J. Probab.* 6 (2001), no. 23, 13 pp. (electronic). ISSN: 1083-6489.

[CAF15] D. Cai, N. Ackerman, and C. Freer. “Priors on exchangeable directed graphs”. In: *ArXiv e-prints* (Oct. 2015). arXiv: 1510.08440 [math.ST].

[CD15] H. Crane and W. Dempsey. “A framework for statistical network modeling”. In: *ArXiv e-prints* (Sept. 2015). arXiv: 1509.08185 [math.ST].

[CF14] F. Caron and E. B. Fox. “Sparse graphs using exchangeable random measures”. In: *ArXiv e-prints* (Jan. 2014). arXiv: 1401.1137 [stat.ME].

[CSKM13] Sung Nok Chiu, Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke. *Stochastic Geometry and Its Applications*. 3rd ed. Wiley, 2013.

[DJ08] Persi Diaconis and Svante Janson. “Graph limits and exchangeable random graphs”. In: *Rendiconti di Matematica, Serie VII* 28 (2008), pp. 33–61. eprint: 0712.2749.

[DVJ03a] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume I: elementary theory and methods*. Second. Springer Science & Business Media, 2003.

[DVJ03b] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*. Second. Springer Science & Business Media, 2003.

[Dur10] R. Durrett. *Probability: Theory and Examples*. 4th Edition. Cambridge U Press, 2010.

[GS69] I. I. Gikhman and A. V. Skorokhod. *Introduction to the Theory of Random Processes*. Dover, 1969.

[Gly87] Peter W. Glynn. "Upper bounds on Poisson tail probabilities". English. In: *Operations Research Letters* 6.1 (1987), pp. 9–14.

[HRH02] Peter D Hoff, Adrian E Raftery, and Mark S Handcock. "Latent space approaches to social network analysis". In: *Journal of the American Statistical Association* 97.460 (2002), pp. 1090–1098.

[HSM15] T. Herlau, M.N. Schmidt, and M. Morup. "Completely random measures for modelling block-structured networks". In: *ArXiv e-prints* (2015). arXiv: 1507.02925.

[Hol12] F. den Hollander. *Probability Theory: The Coupling Method*. 2012.

[Hoo79] D. N. Hoover. *Relations on probability spaces and arrays of random variables*. Tech. rep. Institute of Advanced Study, Princeton, 1979.

[Kal01] O. Kallenberg. *Foundations of Modern Probability*. 2nd. Springer, 2001.

[Kal05] O. Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Springer, 2005.

[Kal90] Olav Kallenberg. "Exchangeable random measures in the plane". English. In: *Journal of Theoretical Probability* 3.1 (1990), pp. 81–136. ISSN: 0894-9840.

[Kal99] Olav Kallenberg. "Multivariate sampling and the estimation problem for exchangeable arrays". In: *J. Theoret. Probab.* 12.3 (1999), pp. 859–883. ISSN: 0894-9840.

[Ker14] D. Kerr. *Ergodic Theory: Independence and Dichotomies*. 2014. URL: http://www.math.tamu.edu/~kerr/book/ (visited on 10/28/2015).

[LOGR12] J. R. Lloyd, P. Orbanz, Z. Ghahramani, and D. M. Roy. "Random function priors for exchangeable arrays". In: *Adv. Neural Inform. Process. Syst. (NIPS) 25*. 2012, pp. 1007–1015.

[LS06] L. Lovász and B. Szegedy. "Limits of dense graph sequences". In: *J. Combin. Theory Ser. B* 96 (2006), pp. 933–957.

[LS07] László Lovász and Balázs Szegedy. "Szemerédi's Lemma for the Analyst". In: *Geom. Func. Anal.* 17.1 (2007), pp. 252–270.

[Lov13] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society, 2013.

[MJG09] Kurt Miller, Michael I Jordan, and Thomas L Griffiths. "Nonparametric latent feature models for link prediction". In: *Advances in neural information processing systems*. 2009, pp. 1276–1284.

[NS01] Krzysztof Nowicki and Tom A. B. Snijders. "Estimation and prediction for stochastic blockstructures". In: *J. Amer. Statist. Assoc.* 96.455 (2001), pp. 1077–1087. ISSN: 0162-1459.

[OR15] P. Orbanz and D.M. Roy. "Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 37.2 (Feb. 2015), pp. 437–461. ISSN: 0162-8828.

[Orb17] P. Orbanz. *Subsampling and invariance in networks*. Preprint. 2017.

[VR15] V. Veitch and D. M. Roy. "The Class of Random Graphs Arising from Exchangeable Random Measures". In: *ArXiv e-prints* (Dec. 2015). arXiv: 1512.03099 [math.ST].

[VR16] V. Veitch and D. M. Roy. "Sampling and Estimation for (Sparse) Exchangeable Graphs". In: *ArXiv e-prints* (Nov. 2016). arXiv: 1611.00843 [math.ST].

[WO13] P. J. Wolfe and S. C. Olhede. "Nonparametric graphon estimation". In: *ArXiv e-prints* (Sept. 2013). arXiv: 1309.5936 [math.ST].

[Jan16] S. Janson. "Graphons and cut metric on sigma-finite measure spaces". In: *ArXiv e-prints* (Aug. 2016). arXiv: 1608.01833 [math.CO].

[Jan17] S. Janson. "On convergence for graphexes". In: *ArXiv e-prints* (Feb. 2017). arXiv: 1702.06389 [math.PR].

# DECLARATION

Put your declaration here.

*Toronto, May 2017*

Victor Veitch

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

*Final Version* as of May 11, 2017 (`classicthesis` version 4.2).