Random Discrete Structures: Approximations and Applications

The investigators in this discovery grant are:

Aihua Xia at http://www.ms.unimelb.edu.au/Personnel/profile.php?PC_id=76
Andrew Barbour at http://user.math.uzh.ch/barbour/
Phil Pollett at http://www.maths.uq.edu.au/~pkp/
Nathan Ross at http://www.ms.unimelb.edu.au/~rossn1/
Malwina Luczak at http://www.maths.qmul.ac.uk/~luczak/
Gesine Reinert at http://www.stats.ox.ac.uk/~reinert/
Adrian Roellin at http://www.stat.nus.edu.sg/~staar/
Erol Pekoz at http://people.bu.edu/pekoz/

Aims and background

Mathematical modelling is the tool of choice for understanding real world systems when empirical experimentation is too expensive or logistically infeasible. For example, in a power grid the failure of one component can have a deleterious effect on adjacent components, potentially causing a cascade of failures [25]. Since determining the chance of a cascade by experimentation is not practical, the wise course is to use data from past failures of components to construct a model which can be used to forecast the chance of a future cascade.

The behaviour of many real world systems can be modelled by random discrete structures (RDS) that evolve over time. For example, a power grid can be modelled as an evolving random network: substations, transformers, and power plants are connected, disconnected and rerouted as demands change. Or, the sizes of populations of frogs in adjacent patches of forest (so-called metapopulations) can be modelled as interdependent stochastic population processes. Important quantities of interest in RDS, such as the chance that a random network becomes disconnected or the chance that the population in a patch of forest goes extinct, cannot always be computed exactly. The standard practice then is to approximate the finite time quantity by an analogous long-run quantity which can be computed exactly. The most well-known example is using the central limit theorem (CLT) to approximate the distribution of the empirical mean of a random sample: as the sample size grows, the fluctuations of the empirical mean about the true population mean become approximately Gaussian. In practice, the Gaussian probabilities are substituted directly for those involving the empirical mean, such as in computations of "p-values", which are ubiquitous in science. An important question in this general framework is: what is the error made in such an approximation? To answer it is crucial, because if the errors are larger than can be tolerated in the application, then the approximation is not appropriate. With this perspective in mind, the project aims to study RDS that arise from application areas including ecology, complex networks, insurance, population genetics, computer science, and statistics. In particular the project will

A1. Determine large-scale statistical features, such as connectivity in a network, which arise from the different random rules that define RDS.

A2. Understand the way these features interact with additional processes that interface with RDS (for example failures in a power grid or the spread of disease within a population network).

A3. Approximate quantities of interest in RDS by long-run quantities, and assess the quality of these approximations by providing explicit error bounds.

A4. Extend several standard techniques in applied probability, including branching process and differential equation approximations, so that they can be applied effectively to RDS, in addition to developing new and broadly applicable techniques of independent interest, such as stochastic fixed point methods.

These aims are critically important in model selection and verification: if some features of a model match empirical data well, then it may be used to predict values of unobservable data or the effects

of changes to the system. Two models used to explain the so-called power law behaviour of the empirical degree distribution (defined to be the histogram of the data of the number of adjacent neighbours of each node of a network) observed in power grids are the preferential attachment model of the seminal paper [8] and the highly optimised tolerance model of [27]. Which model is more appropriate for a given application? Should a third model be used that interpolates between the two approaches [34]? Or, perhaps none of these models is appropriate. Of course addressing this issue relies on the data at hand, the relevant properties of the models, and the agreement between the two. Our purpose is to determine properties of these and other models with the ultimate goal of model selection and verification. Germane to these questions is our past work [64], which is characteristic of the kind of results the project aims to achieve. The approaches of [8, 27, 34] all show that when the number of nodes in their model becomes large, the degree distribution of the network roughly obeys a power law. But other features of the models differ, and these other features will determine the appropriateness of the model in a given application. Thus other aspects of these models need to be studied and understood, and for the preferential attachment model this is the content of [64], where the probabilities corresponding to the number of neighbours of a *given* node are approximated (A1) with an *error* on the approximation (A3). (See also [24, 66, 72] for error bounds on the corresponding approximation for the power law behaviour of the preferential attachment model.) Moreover, the work of [64, 66, 72] required the use of a new stochastic fixed point technique used in conjunction with a novel development of a powerful tool in probability approximation and limit theory called Stein's method (A4).

The project will focus on features of RDS (A1), RDS in time and space (A2), the asymptotic behaviour of RDS (A1, A2), the accuracy of our approximations (A3) and developing novel methods (A4). Each is detailed below.

Features of RDS. A significant focus of the project will be to investigate properties of RDS that arise from real world systems. In the context of random networks, the practitioner has many models to choose from, and each one has different features that make it appropriate or useful in different situations; see [59] and the collection [60] for historical overview and context. The simplest model of a random network is the Erdős-Rényi random network, where each pair of nodes is attached by an edge independently according to the outcome of a (possibly biased) coin toss. These models are reasonably well understood (see the texts [21, 46]), but there are still detailed questions to address [69, 70]. A related model is the stochastic block model, where edges are still connected independently, but the nodes are split into groups $1, \ldots, K$ such that the chance a node in group i connects to a node in group j is given by parameters $p(i, j)$. This is a popular model used for community detection ($p(i, i) > p(i, j)$ for $i \neq j$) [20]. Two models that are widely believed to have realistic features and are close to this proposal are the small worlds model of [77] and the configuration model. Both are used regularly in epidemic modelling; see under the next heading for further explanation. Finally we mention again the preferential attachment model discussed previously. Since the seminal paper [8], there has been an explosion of research activity, with the first rigorous mathematical results due to [22, 23]. As described previously, the work of [64, 66, 72] develops new methods for deriving properties of the degree distributions of these models. Two concrete aims of the project are to extend these methods to apply to generalisations of the basic model, such as those of [17, 30, 73], and to other random network models; and to push them further to derive more properties. For example, what is the *joint* distribution of a given collection of nodes in these models?

Many general RDS can be understood through urn models. In fact, urn models are abstracted from applications in statistics (experiment design in clinical trials) [78], computer science (analysis of algorithms) [52], evolutionary biology [43], and some random networks, including the preferential attachment network [17, 64]. The basic general model is of an urn initially containing black and white balls. At each step in the process, a ball is drawn and returned to the urn, and some number of black and white balls is also added, depending on the color of the ball drawn. These models exhibit a "rich get richer" phenomenon, that appears in many applications. Another feature of the models is that the behaviour of the (random) number of white (or black) balls in the urn after many steps depends very sensitively on the rule for returning balls to the urn; thus new models arising in applications often

have to be understood from scratch. The work of [64] provides approximations with errors for the random number of white balls in families of such urn models, and required the development of new methods. The project will continue the exploration of generalisations and embellishments (different rules for returning balls, more colours, et cetera), especially using the methods of [64].

RDS in time and space. RDS are interesting in their own right. However, they frequently appear in applications as an underlying structure on which a second random process evolves. As an example, the spread of infection in a population, whose contact structure is described by a random network realized from the configuration model [76], has some features more realistic than those of the classical epidemic models that are based on an underlying Erdős–Rényi random graph, and its predictions differ quantitatively from those of the classical models. In HIV models, such as that of [57], the underlying network may itself also evolve in time. The 'gossip' models of [4], describing the spread of information, also have this feature, in that local dissemination is deterministic, but there is also long–range spread resulting from occasional contacts that occur randomly in time and space. An aim of the project is to determine properties and features of such processes that evolve on top of random network structures (such as those mentioned in the previous subsection), again with the goal of guiding practitioners to models appropriate for their particular application.

Asymptotic behaviour of RDS. A number of recent empirical studies indicate that variation in the quality of habitat plays an important role in the distribution pattern of species, and can have a significant impact on the dynamics of the population as a whole, in particular on its viability [58]. Given the considerable resources devoted to managing populations in Australia, for both commercial and conservation purposes, it is vital that the role of patch variation be understood. We aim to develop mathematical tools for studying the dynamics of metapopulations in which patch characteristics are allowed to vary. This aim aligns with those of "RDS in time and space", since positions and qualities of patches in a metapopulation model can also be viewed as the realisation of a random process; in [62], a random field with non-trivial dependence structure is assumed. The metapopulation then evolves as a random process, where the influence between the events at different patches depends on their qualities and relative positions. We will employ branching process approximations to evaluate the chance that a near extinct metapopulation becomes established, and we will develop deterministic and Gaussian approximations to study the growth and long-term equilibrium level of the population, once it has become established.

Accuracy of approximations. A significant effort will be devoted to the study of the accuracy of approximations. Two types of questions will be addressed: (a) distributional approximation errors and (b) extremal properties and tail probability approximation errors.

In relation to (a), the following example is a prototypical case of the range of problems that the project will address. Probability models are used extensively in evolutionary biology and population genetics. While some evolutionary dynamics can be understood using deterministic models, there are other features that are best understood through randomness, for example mutation and selection. Central to population genetics is Kingman's coalescent. This has many wonderful theoretical features, and is also the basis for algorithms used to infer evolutionary history from genetic data [75]. It is the basis for estimating the past size of human populations, leading to an understanding of our species prehistory [51, 74], and for estimating the past size of a disease outbreak, using only present genetic material related to the infection [33]. However, the coalescent is used as a proxy for a discrete object called the Wright-Fisher model, and with ever-increasing sample sizes, arising from the abundance of genetic data from cheaper sequencing technologies, understanding the errors in this and other approximations, routinely made in designing estimation algorithms, is of great importance [19,37,50]. A component of the project is to obtain quantitative bounds on the effects on the inference drawn in evolutionary biology of the errors in such approximations.

Apropos of (b), the extremal events associated with RDS, such as earthquakes, financial crises and floods, are perennial in human history. As such events have a small chance of occurring, we are often unprepared for the damage they cause. Although there is a long history of recorded disasters such as floods and earthquakes, we still do not understand how such events arise, and hence how to predict and efficiently prepare for their occurrence. Moreover, in classical statistics, extreme events

were often treated as outliers, and hence did not get sufficient attention in modelling and forecasting. The systematic mathematical study of extreme events can be traced back to [35, 36]. Since then, the asymptotic behaviour of extreme values has been better studied; see for example [1, 49, 61].

However, most existing studies on extreme events rely on the assumption that the extremes, whether observed or not, happen as independently occurring (rare) events, so that their consequences under various scenarios can be more easily assessed. Prediction and prevention regimes based on such studies often have disastrous consequences. Recent examples include the 2007-2008 global financial crisis and the 2011 Tohoku earthquake and tsunami. Practitioners have largely been unaware of the importance of dependence between extreme events, and the information about this dependence that can be gleaned from existing data. Our aim is to use such data, based on the more frequent occurrences of less extreme events, to fit models that properly allow for dependence, and then to study the extremal properties of RDS, given that at least one such event has taken place. We will derive suitable approximate models for the distribution of extreme events under this conditioning, and identify the conditions under which the approximate models are effective.

Research project

Features of RDS. The core idea in the stream of work [64, 66, 72] is to relate various preferential attachment models to certain single colour Pólya urn models (embellishments of those of [44]) and then to determine new properties and features of these models, by developing and applying a distributional fixed point technique, discussed in "Developing Novel Methods", below. The project will develop this work in several directions. First, the preferential attachment model that we previously studied has the rule that a new vertex attaches to an existing vertex with chance *proportional* to its degree. There are some closely related preferential attachment models [17, 18], and others where the probability of attachment is proportional to some power of the degree [47] or, more generally, to an increasing function of the degree ([73]; see also [30]); for some of these rules, our methods apply, and analogues of our results hold. Secondly, even in the proportional attachment models, little is so far known about the joint degrees of multiple vertices, and the project will extend our methods to this setting; see [67] for preliminary steps in this direction. Significant computer experimentation and simulation are necessary for this, as a rough preliminary guide to what theoretical results are to be expected.

Additionally, the correspondence between random graph attachment rules and urn models is rich, and can be exploited in both directions. Given an urn model, a random graph model can be developed for which the degree of a distinguished node corresponds to the number of balls of a given colour in the urn, and vice versa; keeping track of the joint degrees of multiple nodes corresponds to multi-color urn models. So the techniques and methods used in our proportional attachment work have implications for urn models, including some models with more exotic rules than are usual [65].

We lastly mention that the connection between Pólya urn models and random graphs has found use in showing graph level convergence (in the local weak sense) of certain preferential attachment models [17], and a connection to process level convergence (to Aldous's CRT [2, 3]) is discussed in [65, Remark 2.6]. A long term goal of the project is to push our methods to the process level, and perhaps into the burgeoning field of graph limits.

RDS in time and space. Here, we will focus on two areas of practical interest: epidemic modelling and metapopulations.

For epidemics on the configuration graph, results to date are limited to laws of large numbers [16, 32, 45], and some description of their stochastic behaviour, such as through CLTs, would be welcome. Barbour and Röllin have an approach, based on Stein's method, that may yield a CLT for the final size, though even the computation of the variance poses substantial problems. For gossip models, the situation is somewhat easier, and here one can expect to be able to establish not only a CLT for the final size, but also a description of the evolution of the process in time, in terms of a diffusion process approximation. In both settings, and in others such as the household models of [7], the aim is to supplement any approximations with measures of their accuracy, to underpin their use in practice. Extensive simulation studies will also be conducted, to investigate the stability of the approximations with respect to model mis-specification. We will also analyse a family of HIV models

with contact structure evolving randomly in time, which includes a parameter describing how fast the network evolves relative to the evolution of the epidemic process; this may give useful insight into how well either the fixed network model or the totally random network can be used as a surrogate for more realistic models.

For metapopulations, Barbour, McVinish and Pollett are currently examining how well the commonly used deterministic patch occupancy models of [42] actually approximate their stochastic counterparts. To date, we have shown that the approximation is in general not at all good, unless the evolution in each patch is significantly influenced by a large number of other patches. However, the models so far considered are the simplest of their kind. Our aim now is to examine what happens for the more realistic, structured metapopulation models [41], as well as to refine the conclusions of [62] with some measure of approximation error. Structured metapopulation models are most naturally treated in an infinite dimensional setting, entailing substantial mathematical difficulty. Earlier work of [14, 15] should provide pointers for the techniques required to answer such questions. One goal is to derive 'propagation of chaos' results, linking the behaviour of small groups of patches with that of the ensemble. This is particularly important when studying processes such as adaptive evolution, where a new strategy has to be successful at the individual level, before it can translate to the whole population.

Asymptotic behaviour of RDS. *Branching process approximations* are widely used to study the evolution of an initially small population. In the metapopulation context, one typically assumes that the total number of patches is large, but that initially only few are occupied. The early growth phase is then approximated using a branching process. As the total number of patches increases, the number occupied converges weakly to a branching process (or some variant) on finite time intervals; the chance of total extinction is then easily calculated as the smallest fixed point of a function determined by the distribution of offspring. Under certain conditions, these results can be improved to strong convergence on appropriately chosen time intervals. Although the branching process approximation is most commonly used in the study of epidemics, it is also useful in the study of metapopulations that are close to extinction [26], but will need to be adapted to account for spatial effects and patch variation in a continuum.

Deterministic approximations are widely used to approximate the behaviour of large populations. Typically, this involves the solution to certain (finite dimensional) differential or difference equations. [31] provides a survey of the relevant literature, and gives quantifiable error bounds. Questions concerning population persistence can then be phrased in terms of stability of the equilibrium points of the corresponding differential/difference equation (for which the analysis is typically much simpler). Going beyond the finite dimensional case, [13, 15] studied a Markov population process where each individual has one of countably many types, and applied their results to certain metapopulations and parasitic infections. They were able to prove convergence to a flow on a certain weighted ℓ_1 space and also gave a rate of convergence. Establishing a deterministic limit for a Markov population process incorporating variation of individuals (in this case patches) is a challenging problem. However, there has been significant recent progress in this area for metapopulation models [53, 55, 56]. Our results established only convergence to a deterministic limit. We now propose to construct bounds on the error to assess the quality of the approximation. A refinement of the description of the error beyond a simple bound involves examining if the error, appropriately scaled, converges in some sense. Typically, this involves showing that the scaled error process converges to some Gaussian process (the *Gaussian approximation*). The classic paper [48] gives conditions for the error process to converge to a Gaussian diffusion for simple Markov population processes. In [14], we were able to derive a similar result for population processes with countably many types. In [54] we were able to prove convergence of the scaled error process to a Gaussian process for the model studied in [53].

Accuracy of approximations. The main tool for studying of the quality of our approximations is Stein's method. To explain how we will implement the idea, consider Stein's method for the conditional tail probability approximations of the number of extreme events given that at least m such extreme events have already occurred [38]. Given a stationary sequence of random variables $X_1, \ldots, X_n$ with common distribution function F_X, define the number of exceedances above the threshold s as $N_{s,n} :=$

$\sum_{i=1}^{n} \mathbf{1}_{(s,\infty)}(X_i)$ and the fragility distribution of order m as $FD_{n,m} := \lim_{s \nearrow x_F} \mathcal{L}\left(N_{s,n} | N_{s,n} \geqslant m\right)$, where $x_F := \sup\{t : \ F_X(t) < 1\}$.

We say that a random variable P has a compound Poisson (CP) distribution $\mathrm{CP}(\boldsymbol{\lambda})$ with $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots)$, if $P \stackrel{d}{=} \sum_{i=1}^{\infty} i X_i$, where X_i follows Poisson distribution with mean λ_i and the X_i are independent. One can easily verify that $W \sim \mathcal{L}(P|P \geq 1)$ with $P \sim \mathrm{CP}(\boldsymbol{\lambda})$ if and only if for all bounded functions g on integers,

$$\mathbb{E}\left[\sum_{j=1}^{\infty} j\lambda_j g(W+j) - Wg(W)\mathbf{1}_{\{W>1\}}\right] = 0.$$

This identity provides a characterisation of the conditional CP distribution, and it can be used to study conditional CP approximation errors via Stein's method.

Developing novel methods. The project will strengthen and generalize a number of standard methods in applied probability, to treat problems that are at present beyond their scope. A significant component concerns Stein's method of distributional approximation, and its exploitation using fixed points of distributional transformations; see the survey [71] for an introduction to Stein's method from this point of view.

Stein's method has been used successfully in applications of the kinds to be pursued here: in random networks [28, Chs 4 and 6], [11,64,66,72], computational biology [11, Ch 10], [5,29], branching processes [63], and statistics [6, 39, 40]. However, although we shall use Stein's method in a variety of new applications, our focus is on developing it for use in more difficult situations. For example, the work of [64] studies the random degree of a single vertex in the preferential attachment model; this is a 1-dimensional result. It is of still greater interest to understand the joint distribution of the degrees of a collection of k vertices; a k-dimensional result that contains the 1-dimensional result. Stein's method has been applied in multi-dimensional problems, see for example [68], but much less is known than in the 1-dimensional setting. A main aim of the project is to develop the method for effective use in multi-dimensional applications. A further aim is to develop the method at the *process* or *infinite-dimensional* level. For example, consider the vertex degrees in the preferential attachment model. It is possible to define an abstract object that represents the random network after adding infinitely many vertices and their edges [17]. Is it possible to devise a version of Stein's method that will measure the closeness of the system with n vertices to the limit? This is an extremely challenging problem. However, the investigators have experience in developing the method for settings more complicated than the approximation of random vectors [9, 10, 12, 29], and the goal is to push these techniques even further.

References

[1] S. Albeverio, V. Jentsch, and H. Kantz, *Extreme events in nature and society*, Springer, 2006.

[2] D. J. Aldous, *The continuum random tree. I*, Annals of Probability, 19 (1991), pp. 1–28.

[3] ——, *The continuum random tree. III*, Annals of Probability, 21 (1993), pp. 248–289.

[4] D. J. Aldous, *When knowing early matters: gossip, percolation and Nash equilibria*, in Prokhorov and Contemporary Probability Theory, A. N. Shiryaev, S. R. S. Varadhan, and E. L. Presman, editors, Springer Proceedings in Mathematics & Statistics, Springer Berlin Heidelberg, 2013, pp. 3–27.

[5] R. Arratia, L. Goldstein, and L. Gordon, *Two moments suffice for Poisson approximations: the Chen-Stein method*, Annals of Probability, 17 (1989), pp. 9–25.

[6] R. Arratia, L. Goldstein, and L. Gordon, *Poisson approximation and the Chen-Stein method*, Statistical Science, 5 (1990), pp. 403–434. With comments and a rejoinder by the authors.

[7] F. Ball, D. Sirl, and P. Trapman, *Analysis of a stochastic SIR epidemic on a random network incorporating household structure*, Mathematical Biosciences, 224 (2010), pp. 53–73.

[8] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–512.

[9] A. D. Barbour, *Stein's method and Poisson process convergence*, Journal of Applied Probability, 25A (1988), pp. 175–184.

[10] ——, *Stein's method for diffusion approximations*, Probability Theory and Related Fields, 84 (1990), pp. 297–322.

[11] A. D. Barbour, L. Holst, and S. Janson, *Poisson approximation*, vol. 2 of Oxford Studies in Probability, The Clarendon Press, Oxford, 1992.

[12] A. D. Barbour and S. Janson, *A functional combinatorial central limit theorem*, Electronic Journal of Probability, 14 (2009), paper no. 81, 2352–2370.

[13] A. D. Barbour and M. J. Luczak, *Laws of large numbers for epidemic models with countably many types*, Annals of Applied Probability, 18 (2008), pp. 2208–2238.

[14] ——, *Central limit approximations for Markov population processes with countably many types*, Electronic Journal of Probability, 17 (2012), paper no. 90, 16 pp.

[15] ——, *A law of large numbers approximation for Markov population processes with countably many types*, Probability Theory and Related Fields, 153 (2012), pp. 727–757.

[16] A. D. Barbour and G. Reinert, *Approximating the epidemic curve*, Electronic Journal of Probability, 18 (2013), paper no. 54, 30 pp.

[17] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi, *Asymptotic behavior and distributional limits of preferential attachment graphs*, Annals of Probability, 42 (2014), pp. 1–40.

[18] S. Bhamidi, J. M. Steele, and T. Zaman, *Twitter event networks and the superstar model*. Preprint `http://arxiv.org/abs/1211.3090`, 2012.

[19] A. Bhaskar, A. G. Clark, and Y. S. Song, *Distortion of genealogical properties when the sample is very large*, Proceedings of the National Academy of Sciences USA, 111 (2014), pp. 2385–2390.

[20] P. J. Bickel and A. Chen, *A nonparametric view of network models and Newman–Girvan and other modularities*, Proceedings of the National Academy of Sciences USA, 106 (2009), pp. 21068–21073.

[21] B. Bollobás, *Random graphs*, vol. 73 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, second edition, 2001.

[22] B. Bollobás and O. Riordan, *The diameter of a scale-free random graph*, Combinatorica, 24 (2004), pp. 5–34.

[23] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures & Algorithms, 18 (2001), pp. 279–290.

[24] G. Brightwell and M. Luczak, *Vertices of high degree in the preferential attachment tree*, Electronic Journal of Probability, 17 (2012), paper no. 14, 43 pp.

[25] C. D. Brummitt, P. D. Hines, I. Dobson, C. Moore, and R. M. D'Souza, *Transdisciplinary electric power grid science*, Proceedings of the National Academy of Sciences USA, 110 (2013), p. 12159.

[26] F. M. Buckley and P. K. Pollett, *Limit theorems for discrete-time metapopulation models*, Probability Surveys, 7 (2010), pp. 53–83.

[27] J. M. Carlson and J. Doyle, *Highly optimized tolerance: A mechanism for power laws in designed systems*, Physical Review E, 60 (1999), p. 1412.

[28] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao, *Normal approximation by Stein's method*, Probability and its Applications, Springer, Heidelberg, 2011.

[29] L. H. Y. Chen and A. Xia, *Stein's method, Palm theory and Poisson process approximation*, Annals of Probability, 32 (2004), pp. 2545–2569.

[30] A. Collevecchio, C. Cotar, and M. LiCalzi, *On a preferential attachment and generalized Pólya's urn model*, Annals of Applied Probability, 23 (2013), pp. 1219–1253.

[31] R. W. R. Darling and J. R. Norris, *Differential equation approximations for Markov chains*, Probability Surveys, 5 (2008), pp. 37–79.

[32] L. Decreusefond, J.-S. Dhersin, P. Moyal, and V. C. Tran, *Large graph limit for an SIR process in random network with heterogeneous connectivity*, Annals of Applied Probability, 22 (2012), pp. 541–575.

[33] A. Drummond, A. Rambaut, B. Shapiro, and O. Pybus, *Bayesian coalescent inference of past population dynamics from molecular sequences*, Molecular Biology and Evolution, 22 (2005), pp. 1185–1192.

[34] R. M. D'Souza, C. Borgs, J. T. Chayes, N. Berger, and R. D. Kleinberg, *Emergence of tempered preferential attachment from optimization*, Proceedings National Academy of Sciences USA, 104 (2007), pp. 6112–6117.

[35] R. A. Fisher and L. H. C. Tippett, *Limiting forms of the frequency distribution of the largest or smallest member of a sample*, Proceedings of the Cambridge Philosophical Society, 24 (1928), pp. 180–190.

[36] M. Fréchet, *Sur la loi de probabilité de l'écart maximum*, Annales de la Societe Polonaise de Mathematique, 6 (1927), pp. 93–116.

[37] Y.-X. Fu, *Exact coalescent for the Wright-Fisher model*, Theoretical Population Biology, 69 (2006), pp. 385–394.

[38] H. L. Gan and A. Xia, *Fragility distributions and their approximations*. Preprint `http://arxiv.org/abs/1402.0934`, 2014.

[39] L. Goldstein and G. Reinert, *Stein's method and the zero bias transformation with application to simple random sampling*, Annals of Applied Probability, 7 (1997), pp. 935–952.

[40] ——, *Zero biasing in one and higher dimensions, and applications*, in Stein's method and applications, vol. 5 of Lecture Notes Series Institute of Mathematical Sciences University of Singapore, Singapore University Press, Singapore, 2005, pp. 1–18.

[41] M. Gyllenberg and I. Hanski, *Single-species metapopulation dynamics: a structured model*, Theoretetical Population Biology, 42 (1992), pp. 35–61.

[42] I. Hanski, *A practical model of metapopulation dynamics*, Journal of Animal Ecology, (1994), pp. 151–162.

[43] F. M. Hoppe, *Pólya-like urns and the Ewens' sampling formula*, Journal Mathematical Biology, 20 (1984), pp. 91–94.

[44] S. Janson, *Limit theorems for triangular urn schemes*, Probability Theory and Related Fields, 134 (2006), pp. 417–452.

[45] S. Janson, M. Luczak, and P. Windridge, *Law of large numbers for the SIR epidemic on a random graph with given degrees*. Preprint `http://arxiv.org/abs/1308.5493`, 2013.

[46] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[47] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Connectivity of growing random networks*, Physical Review Letters, 85 (2000), pp. 4629–4632.

[48] T. G. Kurtz, *Limit theorems for sequences of jump Markov processes approximating ordinary differential processes*, Journal of Applied Probability, 8 (1971), pp. 344–356.

[49] M. R. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983.

[50] S. Lessard, *An exact sampling formula for the Wright-Fisher model and a solution to a conjecture about the finite-island model*, Genetics, 177 (2007), pp. 1249–1254.

[51] H. Li and R. Durbin, *Inference of human population history from individual whole-genome sequences*, Nature, 475 (2011), pp. 493–496.

[52] H. M. Mahmoud, *Pólya urn models*, Texts in Statistical Science Series, CRC Press, Boca Raton, FL, 2009.

[53] R. Mcvinish and P. K. Pollett, *Limits of large metapopulations with patch-dependent extinction probabilities*, Advances in Applied Probability, 42 (2010), pp. 1172–1186.

[54] R. McVinish and P. K. Pollett, *A central limit theorem for a discrete-time SIS model with individual variation*, Journal of Applied Probability, 49 (2012), pp. 521–530.

[55] ——, *The deterministic limit of a stochastic logistic model with individual variation*, Mathematical Biosciences, 241 (2013), pp. 109–114.

[56] ——, *The limiting behaviour of a stochastic patch occupancy model*, Journal of Mathematical Biology, 67 (2013), pp. 693–716.

[57] M. Morris and M. Kretzschmar, *Concurrent partnerships and the spread of HIV*, Aids, 11 (1997), pp. 641–648.

[58] A. Mortelliti, G. Amori, and L. Boitani, *The role of habitat quality in fragmented landscapes: a conceptual overview and prospectus for future research*, Oecologia, 163 (2010), pp. 535–547.

[59] M. E. J. Newman, *The structure and function of complex networks*, Society of Industrial and Applied Mathematics Review, 45 (2003), pp. 167–256 (electronic).

[60] M. E. J. Newman, A.-L. Barabási, and D. J. Watts, *The structure and dynamics of networks*, Princeton University Press, 2006.

[61] S. Y. Novak, *Extreme value methods with applications to finance*, vol. 122 of Monographs on Statistics and Applied Probability, CRC Press, Boca Raton, FL, 2012.

[62] O. Ovaskainen and S. J. Cornell, *Asymptotically exact analysis of stochastic metapopulation dynamics with explicit spatial structure*, Theoretical Population Biology, 69 (2006), pp. 13–33.

[63] E. A. Peköz and A. Röllin, *Exponential approximation for the nearly critical Galton-Watson process and occupation times of Markov chains*, Electronic Journal of Probability, 16 (2011), paper no. 51, 1381–1393.

[64] E. A. Peköz, A. Röllin, and N. Ross, *Degree asymptotics with rates for preferential attachment random graphs*, Annals of Applied Probability, 23 (2013), pp. 1188–1218.

[65] ——, *Generalized gamma approximation with rates for urns, walks and trees*. Preprint `http://arxiv.org/abs/1309.4183`, 2013.

[66] ——, *Total variation error bounds for geometric approximation*, Bernoulli, 19 (2013), pp. 610–632.

[67] ——, *Joint degree distributions of preferential attachment random graphs*. Preprint `http://arxiv.org/abs/1402.4686`, 2014.

[68] G. Reinert and A. Röllin, *Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition*, Annals of Probability, 37 (2009), pp. 2150–2173.

[69] O. Riordan and N. Wormald, *The diameter of sparse random graphs*, Combinatorics, Probability and Computing, 19 (2010), pp. 835–926.

[70] A. Röllin and N. Ross, *Local limit theorems via Landau-Kolmogorov inequalities*. Preprint `http://arxiv.org/abs/1011.3100`, to appear in *Bernoulli*, 2012.

[71] N. Ross, *Fundamentals of Stein's method*, Probability Surveys, 8 (2011), pp. 210–293.

[72] ——, *Power laws in preferential attachment graphs and Stein's method for the negative binomial distribution*, Advances in Applied Probability, 45 (2013), pp. 876–893.

[73] A. Rudas, B. Tóth, and B. Valkó, *Random trees and general branching processes*, Random Structures & Algorithms, 31 (2007), pp. 186–202.

[74] S. Sheehan, K. Harris, and Y. S. Song, *Estimating variable effective population sizes from multiple genomes: A sequentially markov conditional sampling distribution approach*, Genetics, 194 (2013), pp. 647–662.

[75] S. Tavaré, *Ancestral inference in population genetics*, in Lectures on probability theory and statistics, vol. 1837 of Lecture Notes in Mathematics, Springer, Berlin, 2004, pp. 1–188.

[76] E. Volz, *SIR dynamics in random networks with heterogeneous connectivity*, Journal of Mathematical Biology, 56 (2008), pp. 293–310.

[77] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442.

[78] L.-X. Zhang, F. Hu, S. H. Cheung, and W. S. Chan, *Immigrated urn models—theoretical properties and applications*, Annals of Statistics, 39 (2011), pp. 643–671.