Graph Matching: Statistical Perspectives and Structural Results

Joshua Agterberg

December 19, 2018

1 Introduction

The graph matching problem has attracted a great deal of attention recently. Indeed, different disciplines such as theoretical computer science, pattern recognition, machine learning, and statistics have all contributed to the literature on graph matching. Graph matching, or the problem of aligning two vertex sets of graphs, has both an *inexact* and an *exact* problem formulation. In the exact graph matching problem, two graphs are assumed to be isomorphic and the goal is to find the true permutation. In the inexact graph matching problem, graphs are not assumed to be isomorphic, and instead a permutation that minimizes disagreements between the two graphs is sought.

The literature on graph matching has developed through several different avenues. Both the inexact and exact graph matching problems have attracted researchers in theoretical computer science to study the complexity and propose algorithmic solutions. For more details on algorithms and "deterministic" studies of graph matching, see the survey papers Conte et al. (2004) and Emmert-Streib et al. (2016). From the statistical perspective, the literature has taken on a bit of a different flavor; typically, a statistical model is introduced and probabilistic statements are made about matchability or consistency of algorithms. Such a formulation also allows for the powerful tools of information theory to be brought to bear.

In this paper, I will discuss problems related to consistency of graph-matching algorithms and matchability of random graphs. In general, I will not focus on specific algorithms. In particular, I will cover results that (to me) reveal the structure of the graph matching problem. I will use the following questions as guidance in determining what to cover:

- How general can models be such that graphs are matchable?
- Given matchability, are there algorithms that can approximate the solution well?
- Given matchability, is there a deeper relationship between finding the true match and computation?

The outline of the paper is as follows: in Section 2, I will give definitions and notation, and in Section 3 I wll give the problem formulation. In Sections 4 and 5I will discuss a few of the important results in matchability and consistency of graph matching. In Section 6 I will cover a few important technical concepts necessary for many of the proofs in the area, and in Section 7 I will conclude and cover a few open problems.

2 Definitions and Notation

A graph $G = (V, E)$ consists of vertices (V) and edges (E). A compact way of representing a graph is via the adjacency matrix $A \in \{0, 1\}^{n\times n}$ where n is the number of vertices. For two matrices C and D, let $C \oplus D = \begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}$ denote their direct sum. I will write 0 as the zero matrix and $J = 11^\top$ the matrix of all ones, with dimension either as a subscript or clear from the context.

To even begin to make probabilistic statements about graph matching, one needs to consider some sort of distribution. To this end, the statistical model we will be considering is the following.

Definition 1 (Correlated Bernoulli). Suppose $\Lambda^1, \Lambda^2 \in [0, 1]^{n\times n}$ and $R \in [0, 1]^{n\times n}$. Two graphs G_1 and G_2 are said to be $R-$ Correlated Bernoulli random graphs if their adjacency matrices (A, B) satisfy

1. B_{uv} are independent with $B_{uv} \sim \text{Bernoulli}(\Lambda^2_{uv})$

2. A_{uv} are independent with $A_{uv} \sim \text{Bernoulli}(\Lambda^1_{uv})$

3. The Pearson correlation $\rho(A_{uv}, B_{uv}) = R_{uv}$.

We write $(A, B) \sim R - \text{CorrBern}(\Lambda^1, \Lambda^2)$.

Note that the above model encompasses a wide range of models; for example, the random dot product graph, stochastic block model, and Erdős-Rényi model are all special cases of the above. See Athreya et al. (2017) for more information on these other families of distributions. However, in many of the papers examined below, it may be the case that $\Lambda^1 = \Lambda^2$, or if A and B are graphs on a different number of vertices, only the induced subgraph of the first n_c vertices are correlated (with $n_c < n$ representing the "core" vertices). In any case, it will be clear below if a distinction needs to be made.

Remark 1. Several authors examine the same model above under a different guise (e.g. Cullina and Kiyavash (2017); Cullina et al. (2018)). When $\Lambda^1 = \Lambda^2 = pJ_n$ and $R = \rho J$ for some $\rho > 0$, the above model is a ρ-correlated Erdős-Rényi random graph. A different, but equivalent formulation is to first generate a parent graph $G_0 \sim ER(p)$ and then to sample edges from G_0 with some probability s independently to generate G_1 and G_2. Both models can be considered reparamaterizations of one another.

3 Problem Formulation

For two graphs G_1 and G_2 on n vertices, the graph matching problem is to find an element of

$$\arg\min_{P \in \mathcal{P}} ||A - PBP^\top||_F^2$$

where $\mathcal{P}$ is the set of permutation matrices. Note by expanding the product out, this is equivalent to solving

$$\underset{P \in \mathcal{P}}{\text{argmax}}\, \text{trace}\left(APBP^T\right).$$

In some cases, it may be case that G_2 is a graph on $m \neq n$ vertices, in which case the problem will have to be modified. In addition, someauthors consider graph matching with the use of *seed vertices*; i.e. vertices in which the correspondence is already known. In this case, the graph matching problem solution is assumed to be of the form $I_k \oplus P_{n-k}$ where k is the number of seeds. I will not be emphasizing seeded graph matching.

One natural question to ask is whether graph matching is even possible. To this end, I introduce the notion of *matchability*, taken from Lyzinski and Sussman (2017).

Definition 2. We say that $(A, B) \sim R - \text{CorrBern}\left(\Lambda^1, \Lambda^2\right)$ are matchable if

$$\arg\min_{P \in \mathcal{P}} \|A - PBP^\top\|_F = \{I_n\},$$

where I_n is the $n \times n$ identity matrix.

The above definition highlights the fact that the minimizer of the objective function may not be the identity. In this case, even though there is a latent correspondence, minimizing the objective function (even exactly), will not find the true solution.

4 Matchability

The purpose of this section is to address the question:

- How general can models be such that graphs are matchable?

I present results according to model complexity, focusing on the question of matchability.

4.1 Erdős-Rényi Model

From our general model above, it could be quite feasible that two graphs are not matchable; i.e. there is some permutation P such that P minimizes $||A - PBP^\top||_F$ by more than the identity. Hence, one way of tackling this problem would be to consider whether graphs are matchable with high probability. One of the first results on this problem is the pioneering work in Wright (1971) and, more recently, those of Cullina and Kiyavash (2017) in which conditions are derived for which exact matching is possible under a joint Erdős-Rényi model (i.e. $\Lambda^1 = r_1 J$, $\Lambda^2 = r_2 J$ for some r_1 and r_2, and $R = \rho J$.) One of their main results is given below (under an equivalent model formulation specifying the joint distribution rather than the marginals).

Theorem 1 (Theorem 3 From Cullina and Kiyavash (2017))**.** If the joint distribution satisfies satisfies

$$\left(\sqrt{p_{11}p_{00}} - \sqrt{p_{10}p_{01}}\right)^2 \geq \frac{2\log n + \omega(1)}{n},$$
$$\frac{p_{01}p_{10}}{p_{11}p_{00}} < 1,$$

then both graphs are matchable with probability $1 - o(1)$. Here $p_{ij} = P(A_{uv} = i, B_{uv} = j)$ is the joint distribution for (A_{uv}, B_{uv}).

The authors also provide a converse direction to prove non-matchability, but I will leave it out for brevity. Note that the above result is making an assumption about the correlation; the theorem says that if the correlation is sufficiently strong between both graphs, then with high probability both graphs are matchable. This will be a theme in several of the results to follow.

In a similar spirit to Cullina and Kiyavash (2017), the authors of Cullina et al. (2018) derive information theoretic limits for partial recovery of random graphs. Their main theorem states that with probability tending to one, an increasingly large fraction of vertices can be correctly matched – the "core" vertices. A converse is also provided – the main technical difference between Cullina and Kiyavash (2017) andCullina et al. (2018) is that in Cullina et al. (2018) the average degree need only grow with n, and need not grow a a logarithmic rate with n.

4.2 Stochastic Block Models

Since Erdős-Rényi models have been shown to give sub-par performance of real graphs in practice, extending the results above to more general models has been important. In particular, both Shirani et al. (2018) and Onaran et al. (2016) provide theorems in which graphs are matchable under the ρ-correlated Stochastic Block Model. However, one of the deeper results is that of Lyzinski (2016) in which analogous theorems to Cullina and Kiyavash (2017) are derived under the ρ-correlated stochastic block model. In this model, edge probabilities are determined by a $K \times K$ matrix Λ and a block assignment vector b such that $\Lambda^1_{uv} = \Lambda^2_{uv} = \Lambda_{b_u b_v}$. I restate the matchability result below.

Theorem 2 (Theorem 10 From Lyzinski (2016))**.** Let A and B be the adjacency matrices of $\rho - SBM(K, \vec{n}, b, \Lambda)$ random graphs with K and Λ fixed in n . There exists a constant $\alpha > 0$ such that if $\rho \geq \sqrt{\alpha \log n / n}$, then

$$\mathbb{P}\left(\exists P \in \Pi(n) \backslash \{I_n\} \text{ s.t. } \left\|A - PBP^T\right\|_F < \|A - B\|_F\right) = O\left(e^{-3\log n}\right).$$

In addition, Lyzinski (2016) gives a converse result and examines the effect of information loss on subsequent inference. Note that this result generalizes the model in the previous subsection, but does not quite have the same level of generality as the correlated Bernoulli model. In particular, when $K = 1$, this theorem reduces to that of Cullina and Kiyavash (2017) (after suitable reparameterization). The statement above says that with high probability, the two graphs are matchable provided that $\rho = \omega(\sqrt{\log n / n})$ and the probability matrix is fixed in n.

4.3 Bernoulli Model

In all the above models, conditions for matchability are given assuming Erdős-Rényi or the Stochastic Block Model. Generalizing even further, in Lyzinski and Sussman (2017) the authors show that the centering with universal singular value thresholding (Chatterjee, 2015) gives matchability with high probability in a general correlated Bernoulli model with some junk vertices. They develop this via first showing oracle centering gives matchability with high probability (under no assumptions on the graphs), and then show that USVT estimates the the oracle centering sufficiently well by assuming a low-rank structure for Λ^1 and Λ^2 and logarithmic expected degree (which is needed for the USVT estimator to be consistent). The most similar theorem to those above is Theorem 5 from Lyzinski and Sussman (2017), restated below.

Theorem 3 (Theorem 5 From Lyzinski and Sussman (2017))**.** Let $(A, B) \sim R - corrBern(\Lambda^1, \Lambda^2)$ and consider $\tilde{A} := A - \Lambda^1$ and $\tilde{B} := B - \Lambda^2$. Define

$$\epsilon = \min_{u,v} R_{uv} \sqrt{\Lambda^1_{uv} \left[1 - \Lambda^1_{uv}\right] \Lambda^2_{uv} \left[1 - \Lambda^2_{uv}\right]}.$$

If $\epsilon = \omega(\sqrt{\log(n)/n})$, then $\mathbb{P}(\tilde{A}$ and $\tilde{B}$ are matchable$) \geq 1 - \left\{-\Omega\left(\epsilon^2 n\right)\right\}$.

Note the similarities of the above theorem to previous theorems. In particular, if the correlation is high enough, then both graphs are matchable – in particular, here the graphs are centered to be of zero-mean. Much of the rest of Lyzinski and Sussman (2017) involves finding a way to estimate the centering scheme, since centering assuming knowledge of Λ^1 and Λ^2 would be infeasible in practice.

In addition, much in a similar fashion to Cullina et al. (2018), the authors of Lyzinski and Sussman (2017) show that centering via USVT gives core-matchability provided the core size is sufficiently large.

4.4 $n \neq m$

Although not strictly related to results above, when graphs are different sizes, the theory is not as well developed, so I thought it important to cover a related result. In the recent paper Sussman et al. (2018), the authors examine the graph matching problem under three different scenarios, dubbed "paddings." Assuming that G_1 is a graph on n vertices and G_2 is a graph on $m > n$ vertices, a natural question is, given the existence of a latent correspondence, how does one formulate the graph matching problem? Under the general R-correlated Bernoulli model where $\Lambda^1[n,n] = \Lambda^2 = \Lambda$, where $\Lambda^1[n,n]$ is the $n \times n$ induced subgraph of the first n vertices, the authors consider the following paddings:

- Naive Padding: Letting $\tilde{A} = A \oplus 0_{m-n}$ and match $\tilde{A}$ and B.
- Centered Padding: Let $\tilde{A} := (2A - J) \oplus 0_{m-n}$, $\tilde{B} = 2B - J$, and match $\tilde{A}$ and $\tilde{B}$.
- Oracle Padding: Let $\tilde{A} = (A - \Lambda_n) \oplus 0$, $\tilde{B} = B - \Lambda$ and match $\tilde{A}$ and $\tilde{B}$.

In particular, they prove that the naive padding fails with high probability, and the oracle and centered padding schemes succeed with high probability (with requirements on the correlation matrix for the centered padding scheme). Such a result shows that that naive padding searches for the best *subgraph* as opposed to the best *induced subgraph*.

5 Consistency

Now that matchability has been established for a number of problems, a follow on question is whether there exist consistent algorithms to find the matching given that one exists. Such results is more practical, since finding a consistent algorithm allows one to perform graph matching on real data. This section will focus on the other two questions:

- Given matchability, are there algorithms that can approximate the solution well?
- Given matchability, is there a deeper relationship between finding the true match and computation?

Clearly, the answer to both questions is "yes." For example, an exact algorithm always exists for any graph matching problem– one could simply enumerate every possible permutation. However, such an algorithm would run in $O(n!)$ time, rendering it computationally infeasible. Hence, we wish to seek approximate algorithms – relaxations of the graph matching problem that can find decent solutions. From a theoretical computer science point of view, it may be important to ask whether there exists a polynomial time algorithm to match such graphs given that they are matchable. Indeed, in Mossel and Xu (2018), the authors do just that. However, as existence of polynomial-time algorithms is not really a statistical perspective, I will leave this part of the literature out.

There are generally two ways to reduce the time it takes to run graph matching: 1.) via the use of seed vertices, and 2.) via optimization techniques (e.g. convex relaxation). For the purposes of the question above, having seeds is less of a "structural" result, and more something that either a practitioner has access to or doesn't. Hence, below, I will look at the structural effects different different optimization procedures have on the result of graph matching.

In Lyzinski et al. (2016), the authors consider two different relaxations of the graph matching problem. Since minimizing

$$||A - PBP^\top||_F^2$$

is equivalent to maximizing

$$\text{trace}\left(APBP^T\right),$$

the authors consider two different relaxations. The first problem is convex if P is allowed to range over the doubly-stochastic matrices (its convex hull), and the second problem is indefinite when allowing P to range over its convex hull. However, in Lyzinski et al. (2016), the authors prove that when $R = \rho J$ and $\Lambda^1 = \Lambda^2$, the convex problem almost always fails to find the true permutation, and the the second problem, while infeasible, almost always finds the true permutation if solved exactly. The theorem is reproduced below.

Theorem 4 (Theorem 1 From Lyzinski et al. (2016))**.** Suppose A and B are adjacency matrices for $\rho-$ correlated Bernoulli($\Lambda^1 = \Lambda^2 = \Lambda$) graphs, and there is an $\alpha \in (0, 1/2)$ such that $\Lambda_{i,j} \in [\alpha, 1-\alpha]$ for all $i \neq j$. Let $P^* \in \Pi$, and denote $A' := P^* A P^{*T}$.

1. If $(1-\alpha)(1-\rho) < 1/2$ then it almost always holds that

$$\arg\min_{D\in\mathcal{D}} -\langle A'D, DB\rangle = \arg\min_{P\in\Pi} \left\|A' - PBP^T\right\|_F = \{P^*\}.$$

2. If the between graph correlation $\rho < 1$ then it almost always holds that

$$P^* \notin \arg\min_{D\in\mathcal{D}} \|A'D - DB\|_F.$$

The above result is indeed structural: since efficient algorithms for graph matching are difficult to find in general (unless $P = NP$), finding an approximate algorithm is important when one wishes to implement graph matching in practice. However, as Lyzinski et al. (2016) says, an approximate algorithm can only be consistent when restricted to a computationally infeasible version of the problem!

Such a result might lead one to question whether there is another way to examine the runtime of certain algorithms. This question is partially answered in Fishkind et al. (2018), in which the authors develop a different measure of correlation, the *total correlation* ρ_T and demonstrate empirically that exact graph matching runtime and matchability are both functions of ρ_T. In particular, if μ is the average of all the entries of Λ, and σ^2 is the variance of all such entries, and if $\rho_h := \frac{\sigma^2}{\mu(1-\mu)}$, then ρ_T satisfies $1 - \rho_T = (1-\rho_c)(1-\rho_h)$, where ρ_c is the usual pearson correlation coefficient between the graph. In other words, for a correlated Erdős-Rényi model with the same parameter, $\rho_h = 0$, so $\rho_T = \rho_c$. Intuitively, ρ_T can be viewed as a decomposition of the within-graph correlation and variation (ρ_h), and the between-graph correlation (ρ_c).

However, the in Fishkind et al. (2018), the authors prove strong consistency for an estimator of ρ_T, but do not prove anything about how precisely ρ_T affects graph matching runtime. Perhaps taking advantage of an estimate of ρ_T would allow one to offset the negative results of Lyzinski et al. (2016)?

Regardless, both of the above results show that indeed, there are algorithms that can approximately find the solution, but perhaps not particularly efficiently. In addition, the empirical findings of Fishkind et al. (2018) suggest that ρ_T may provide a deeper, structural connection to finding the true match and the amount of computation time.

6 Important Techniques

In general, the techniques used for the proofs above are not particularly difficult, but instead are quite clever and involve a combination of undergraduate probability and combinatorics. That being said, there are a few specific, clever tactics used a couple times in proving different results, as well as one fairly deep theorem that is implicitly used a few times.

Overall, the results in many of the papers consist of finding clever ways to bound the number of permutations by the number of transpositions that occur and then using observations below. In some cases, moment generating function bounds are used (e.g. Cullina and Kiyavash (2017)), and a union bound is applied over permutations.

6.1 Bivariate Bernoulli Distributions

In Lyzinski and Sussman (2017) the authors note that if X and Y are correlated Bernoulli random variables, then they can be generated by three independent Bernoulli Random variables . In other words if X and Y satisfy

$$(X, Y) \sim \text{BiBern} \begin{pmatrix} p(p + \varrho(1-p)) & p(1-\varrho)(1-p) \\ p(1-\varrho)(1-p) & (1-p)((1-p) + \varrho p) \end{pmatrix},$$

then the Bernoulli variables

$$Z_0 \sim \text{Bern}(p), Z_1 \sim \text{Bern}(p(1-\varrho)), \text{ and } Z_2 \sim \text{Bern}(p + \varrho(1-p))$$

can realize X and Y by simply generating Z_0, and then conditional on Z_0, generate a Bernoulli from either Z_1 or Z_2. For example, $P(X, Y = (1,1)) = p(p + \rho(1-p)) = P(Z_0 = 1)P(Z_1 = 1)$. The authors then use this fact to reduce problems on two correlated Bernoulli random variables to three independent Bernoulli random variables.

6.2 Concentration Inequalities

In many of the papers discussed, concentration inequalities are used as important steps on the way to the proof. Below, I state two such inequalities that get used a number of times in the proofs.

6.2.1 Hoeffding's Inequality

Suppose X_i satisfy $a_i \leq X_i \leq b_i$ almost surely. Then if $S_n := \sum_{i=1}^n X_i$, we have that

$$\mathbb{P}\left(S_n - \mathrm{E}\left[S_n\right] \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n \left(b_i - a_i\right)^2}\right).$$

In the case of Bernoulli random variables (as in many of the proofs in the problems above), the above bound reduces to $\exp(-\frac{2t^2}{n})$. Also, replacing S_n by $\bar{X}$ reduces the bound to e^{-n2t^2}.

The proof of this result comes from Hoeffding's Lemma, which is a bound on the moment generating function, applying Markov's inequality, and minimizing the right hand side over the variable in the moment generating function.

6.2.2 McDiarmid's Inequality

Suppose $X_1, ..., X_n$ are independent and let

$$M := \sup_i \sup_{X_1, ..., X_n} |f(X_1, ..., X_n) - f(X_1, X_2, ..., X_{i-1}, 1 - X_i, X_{i+1}, ..., X_n)|.$$

Let $\sigma^2 := M^2 \sum_i p_i(1-p_i)$ and let $Y = f(X_1, ..., X_n)$. Then $\mathbb{P}(|Y - \mathbb{E}Y| \geq t\sigma) \leq 2e^{-t^2/4}$ for all $0 < t < 2\sigma/M$.

Note that this proof comes from martingale theory; in particular, it uses the concept of a Doob Martingale. Knowing the full intricacies of stochastic processes would not be necessary to apply the inequality, but it is useful if one wants to use a different martingale or find a way to adapt the inequality.

6.3 Birkhoff-Von Neumann Theorem

The Birkhoff-Von Neumann Theorem is a nice theorem that states that the convex hull of the permutation matrices are the doubly-stochastic matrices. I will give a proof sketch here.

First, consider the matrices as vectors in $\mathbb{R}^{n^2}$ with the lexicographic ordering. Then note that the doubly-stochastic matrices satisfy

$$\begin{aligned}
\sum_{i=1}^{n} x_{ij} &\leq 1 \qquad \forall 1 \leq j \leq n \\
-\sum_{i=1}^{n} x_{ij} &\leq -1 \quad \forall 1 \leq j \leq n \\
\sum_{j=1}^{n} x_{ij} &\leq 1 \qquad \forall 1 \leq i \leq n \\
-\sum_{j=1}^{n} x_{ij} &\leq -1 \quad \forall 1 \leq i \leq n \\
x_{ij} &\geq 0 \qquad \forall 1 \leq i, j \leq n.
\end{aligned}$$

The proof proceeds by showing that any vertex of the polyhedron above is a permutation matrix by showing that since there must be n^2 active constraints, exactly $2n$ of them come from the form $\sum x_{ij} = 1$ (one for each i and j) and the other $n^2 - 2n$ are of the form $x_{ij} = 0$ (for all the other i and j).

7 Conclusion

Overall, I have attempted to address partial solutions to the questions

- How general can models be such that graphs are matchable?
- Given matchability, are there algorithms that can approximate the solution well?
- Given matchability, is there a deeper relationship between finding the true match and computation?

For the first question, a common theme was that if the correlation was sufficiently positive between graphs, then graphs could be matchable. In particular, in Lyzinski and Sussman (2017), the authors showed that models can be extremely general assuming oracle centering – however infeasible in practice. Such a result is structural in that just about every moderately correlated random graph model can be matched, but it still paves the way for more structural results. For example, is there a centering that can achieve better quality not assuming the USVT centering scheme? Such a problem amounts to better estimating the probability generating matrix Λ, which is a problem in random graph inference.

For the second and third questions, I discussed results related to computational relaxation of the graph matching problem, which is computationally infeasible in practice. Open questions here amount to finding algorithms that can approximate the solution even better, perhaps using estimates of ρ_T to add some information.

Indeed, there are many open structural questions in graph matching. While I have focused on the question of matchability, similar questions could be asked about *non*-matchability. For example, in Lyzinski (2016), the author conjectures that if $\rho = o(\sqrt{\log(n)/n})$ in the general model, then matching is not achievable, establishing a duality between matchability and non-matchability. Indeed, such a converse is proven in Cullina and Kiyavash (2017). In some of other the papers discussed, theorems are provided in which graph matching is not possible, but it is still an open problem in general.

Overall, graph matching is an exciting field, with intellectual challenges for those in a wide array of disciplines. For the most part, I have stuck to problems from the statistical perspective (hence the title), but there are many questions that could be asked from a theoretical computer science, machine learning, or optimization point of view, such as finding faster or polynomial-time algorithms to approximately solve the graph matching problem. Also, there are a great number of applications, such as for vertex nomination (Patsolic et al., 2017) or database matching. I hope that this note has begun to illuminate some of the statistical and structural issues one might need to consider for graph matching.

References

A. Athreya, D. E. Fishkind, K. Levin, V. Lyzinski, Y. Park, Y. Qin, D. L. Sussman, M. Tang, J. T. Vogelstein, and C. E. Priebe. Statistical inference on random dot product graphs: a survey. *ArXiv e-prints*, September 2017.

Sourav Chatterjee. Matrix estimation by universal singular value thresholding. *Ann. Statist.*, 43(1): 177–214, 02 2015. doi: 10.1214/14-AOS1272. URL `https://doi.org/10.1214/14-AOS1272`.

Donatello Conte, Pasquale Foggia, Carlo Sansone, and Mario Vento. Thirty years of graph matching in pattern recognition. *IJPRAI*, 18:265–298, 05 2004. doi: 10.1142/S0218001404003228.

Daniel Cullina and Negar Kiyavash. Exact alignment recovery for correlated Erd{o}s-Rényi graphs. *arXiv e-prints*, art. arXiv:1711.06783, November 2017.

Daniel Cullina, Negar Kiyavash, Prateek Mittal, and H. Vincent Poor. Partial Recovery of Erd{o}s-Rényi Graph Alignment via k-Core Alignment. *arXiv e-prints*, art. arXiv:1809.03553, September 2018.

Frank Emmert-Streib, Matthias Dehmer, and Yongtang Shi. Fifty years of graph matching, network alignment and network comparison. *Inf. Sci.*, 346(C):180–197, June 2016. ISSN 0020-0255. doi: 10.1016/j.ins.2016.01.074. URL `https://doi.org/10.1016/j.ins.2016.01.074`.

Donniell E. Fishkind, Lingyao Meng, Ao Sun, Carey E. Priebe, and Vince Lyzinski. Alignment Strength and Correlation for Graphs. *arXiv e-prints*, art. arXiv:1808.08502, August 2018.

V. Lyzinski, D. E. Fishkind, M. Fiori, J. T. Vogelstein, C. E. Priebe, and G. Sapiro. Graph matching: Relax at your own risk. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1): 60–73, Jan 2016. ISSN 0162-8828. doi: 10.1109/TPAMI.2015.2424894.

Vince Lyzinski. Information Recovery in Shuffled Graphs via Graph Matching. *arXiv e-prints*, art. arXiv:1605.02315, May 2016.

Vince Lyzinski and Daniel L. Sussman. Matchability of heterogeneous networks pairs. *arXiv e-prints*, art. arXiv:1705.02294, May 2017.

Elchanan Mossel and Jiaming Xu. Seeded Graph Matching via Large Neighborhood Statistics. *arXiv e-prints*, art. arXiv:1807.10262, July 2018.

Efe Onaran, Siddharth Garg, and Elza Erkip. Optimal De-Anonymization in Random Graphs with Community Structure. *arXiv e-prints*, art. arXiv:1602.01409, February 2016.

Heather G. Patsolic, Youngser Park, Vince Lyzinski, and Carey E. Priebe. Vertex Nomination Via Local Neighborhood Matching. *arXiv e-prints*, art. arXiv:1705.00674, May 2017.

Farhad Shirani, Siddharth Garg, and Elza Erkip. Matching graphs with community structure: A concentration of measure approach. *CoRR*, abs/1810.13347, 2018. URL `http://arxiv.org/abs/1810.13347`.

Daniel L. Sussman, Vince Lyzinski, Youngser Park, and Carey E. Priebe. Matched Filters for Noisy Induced Subgraph Detection. *arXiv e-prints*, art. arXiv:1803.02423, March 2018.

E. M. Wright. Graphs on unlabelled nodes with a given number of edges. *Acta Math.*, 126:1–9, 1971. doi: 10.1007/BF02392023. URL `https://doi.org/10.1007/BF02392023`.