From quasirandom graphs to graph limits and graphlets

Fan Chung*
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093
fan@ucsd.edu

Abstract

We generalize the notion of quasirandomness which concerns a class of equivalent properties that random graphs satisfy. We show that the convergence of a graph sequence under the spectral distance is equivalent to the convergence using the (normalized) cut distance. The resulting graph limit is called graphlets. We then consider several families of graphlets and, in particular, we characterize quasirandom graphlets with low ranks for both dense and sparse graphs. For example, we show that a graph sequence G_n, for $n = 1, 2, \ldots$, converges to a graphlets of rank 2, (i.e.,all normalized eigenvalues G_n converge to 0 except for two eigenvalues converging to 1 and $\rho > 0$) if and only if the graphlets is the union of 2 quasirandom graphlets.

1 Introduction

The study of graph limits originated from quasi-randomness of graphs which concerns large equivalent families of graph properties that random graphs satisfy. Lovász and Sós [37] first considered a generalized notion of quasi-randomness as the limits of graph sequences. Since then, there have been a great deal of developments [1, 3, 6, 9, 10, 24, 25, 26, 27, 32, 33, 36, 37, 38, 39, 40, 41, 42, 43] on the topic of graph limits. There are two very distinct approaches. The study of graph limits for dense graphs is entirely different from that for sparse graphs. By dense graphs, we mean graphs on n vertices with cn^2 edges for some constant c. For a graph sequence of dense graphs, the graph limit is formed by taking the limit of the adjacency matrices with entries of each matrix associated with squares of equal sizes which partition $[0, 1] \times [0, 1]$ (see [38, 39]). Along this line of approach, the graph limit of a sparse graph sequence converges to

*Research supported in part by ONR MURI N000140810747, and AFSOR AF/SUB 552082.

zero. Consequently, very different approaches were developed for graph limits of very sparse graphs, mostly with vertex degrees bounded above by a constant independent of the size of the graph [3, 8, 26].

To distinguish from earlier definitions for graph limits (called, graphons, graphines, etc.), we will call the graph limits in this paper by the name of *graphlets* to emphasize the spectral connection. In the subsequent sections, we will give a detailed definition for graphlets as the graph limits of given graph sequences. Although the terminology is sometimes similar to that in differential geometry, the definitions are along the line of spectral graph theory [13] and mostly discrete. In addition, the orthogonal basis of the graphlets of a graph sequence can be used, with additional scaling parameters, to provide a universal basis for all graphs in the domain (or the union of domains) that we consider. In this regard, graphlets play a similar role as the wavelets do for affine spaces.

To study the convergence of a graph sequence, various different metrics come into play for comparing two graphs. For two given graphs, there are many different ways to define some notion of distance between them. Usually the labeling map assigns consecutive integers to the vertices of a graph which can then be associated with equal intervals which partition $[0, 1]$. As opposed to the definitions in previous work, we will not use the usual measure or metric on the interval $[0, 1]$. Instead, our measure on $[0, 1]$ will be determined by the graph sequences that we consider. Before we proceed to examine the distance between two graphs, we remark that there is a great deal of work on distances between manifolds [4, 31] via isometric embeddings. Although the details are obviously different, there are similarities in the efforts for identifying the global structures of the objects of interest. We are using elements of $[0, 1]$ as labels for the (blow-up) vertices, similar to the exchangeable probabilistic measures that were used in [1, 24, 25, 28, 29, 34].

Several metrics for defining distances between two graphs originated from the quasi-random class of graphs [16, 19]. One such example involves the *subgraph counts*, concerning the number of induced (or not necessarily induced) subgraphs of G that are isomorphic to a specified graph F. Another such metric is called the *cut metric* which came from discrepancy inequalities for graphs. The usual discrepancy inequalities in a graph G concern approximating the number of edges between two given subsets of vertices by the expected values as in a random graph and therefore such discrepancy inequalities can be regarded as estimates for the distance of a graph to a random graph. For dense graphs, the equivalence of convergence under the subgraph-count metric and the cut metric among others are well understood (see [9, 38]). The methods for dealing with dense graph limits have not been effective so far for dealing with sparse graphs. A different separate set of metrics has been developed [8, 26] using local structures in the neighborhood of each vertex. Instead of subgraph counts, the associated metric concerns counting trees and local structures in the "balls" around each vertex. The problems of graph limits for sparse graphs are inherently harder as shown in [8]. Nevertheless, most real world complex networks are sparse graphs and the study of graph limits for sparse graphs can be useful for understanding the dynamics of large information networks.

The paper is organized as follows: In Section 2, we first examine the convergence of degree distributions of graphs and we consider the convergence the discrete Laplace operators under the spectral norm. Then we give the definition for graphlets in Section 2.4. In Section 3, we give several families of examples, including dense graphlets, quasi-random graphlets, bipartite quasi-random graphlets and graphlets of bounded rank. In Section 4, we consider the discrepancy distance between two graphs which can be viewed as a normalization of the cut distance. Then we prove the equivalence of the spectral distance and the normalized cut distance for both dense and sparse graphs. Note that our definition of the discrepancy distance is different from the cut distance as used in [8] where a negative result about a similar equivalence was given. In Sections 5 and 6, we further examine quasi-random graphlets and bipartite quasi-random graphlets for graph sequences with general degree distributions. In Sections 7 and 8, we give a number of equivalent properties for certain graphlets of rank 2 and for general k. In Section 9, we briefly discuss connections between the discrete and continuous, further applications in finding communities in large graphs and possible future work that this paper might lead to.

We remark that the work here is different from the spectral approach of graph limits which focuses on the spectrum of the limit of the adjacency matrices in [43]. If the graph limit is derived from a graph sequence which consists of dense and almost regular graphs, the two spectra are essentially the same (differ only by a scaling factor). However, a subgraph of a regular graph is not necessarily regular. All theorems in this paper hold for general graph sequences for both dense graphs and sparse graphs. Some of the methods here can be generalized to weighted directed graphs which will not be discussed in this paper.

2 The spectral norm and spectral distance

For a weighted graph $G = (V, E)$ with vertex set V and edge set E, we denote the adjacency matrix by A_G with rows and columns indexed by vertices in V. For an edge $\{u, v\} \in E$, the edge weight is denoted by $A_G(u, v)$. For a vertex v in $V(G)$, the degree of v is $d_G(v) = \sum_u A_G(u, v)$. We let D_G denote the diagonal matrix with $D_G(v, v) = d_G(v)$. Here we consider graphs without isolated vertices. Therefore, we have $d_G(v) > 0$ for every v and D_G^{-1} is well defined.

We consider the family of operators $\mathcal{W}$ consisting of $W : [0, 1]\times[0, 1] \to [0, 1]$ satisfying , $W(x, y) = W(y, x)$. W is said to be of finite type if there is a finite partition $(S_1, ..., S_n)$ of $[0, 1]$ such that W is constant on each set $S_i \times S_j$. Given a graph G_n on n vertices, a special finite-type associated with G_n is defined by partitioning $[0, 1]$ into n intervals of length $1/n$ and, for a map $\eta : [0, 1] \to V$, the pre-image of each vertex v corresponds to a interval $I_v = (j/n, (j+1)/n]$ for some j. We can define $W_{G_n} \in \mathcal{W}$ by setting :

$$W_{G_n}(x, y) = A_{G_n}(u, v) \tag{1}$$

if $x \in I_u$, and $y \in I_v$.

Suppose we have a sequence of graphs, G_n, for $n = 1, 2, \ldots$. Our goal is to describe the limit of a graph sequence provided it converges. One typical way, as seen in [38], is to take the limit of W_{G_n} under the cut norm. For example, if G_n is in the family of random graphs with edge density $1/2$, the limit of W_{G_n} has all entries $1/2$. However, if we consider sparse graphs such as cycles, then the limit of W_{G_n} converges to the 0 function.

Instead, we will define the graph limit to be associated with a measure space Ω as the limit of measure spaces defined on G_n and the measure μ for Ω is the limit of the measures μ_n associated with G_n. Before we give the detailed definitions of Ω and μ, there are a number of technical issues in need of clarification. The following remarks can be regarded as a companion for the definitions to be given in Sections 2.1 to 2.3 so that possible misinterpretations could be avoided.

Remark 1. We label elements of Ω by $[0, 1]$. However, the geometric structure of Ω can be quite different from the interval $[0, 1]$. In general, Ω can be some complicated compact space. For example, if the G_n are square grids (as cartesian products of two paths), then a natural choice for Ω is a unit square. We will write $V(\Omega) = [0, 1]$ to denote the set of "labels" for Ω while Ω can have natural descriptions other than $[0, 1]$.

Remark 2. In this paper, we mainly concern operators W that are *exchangeable* (see [1, 24, 25, 28, 29, 34]). Namely, for a Lebesgue measure-preserving bijection $\tau : [0, 1] \to [0, 1]$, a rearrangement of W, denoted by W_τ, acts on functions f defined on $[0, 1]$ satisfying

$$Wf(x) = W_\tau f(\tau(x)). \tag{2}$$

We say W is equivalent to W_τ and we write $W \sim W_\tau$. By an exchangeable operator W, we mean the equivalence class of operators W_τ where τ ranges over all measure-preserving bijections on $[0, 1]$.

Remark 3. We consider a family of exchangeable self-adjoint operators $\mathcal{W}^*$ which act on the space of functions $f : [0, 1] \to \mathbb{R}$. Clearly, any exchangeable $W : [0, 1] \times [0, 1] \to [0, 1]$ with $W(x, y) = W(y, x)$ is contained in $\mathcal{W}^*$. The disadvantage of using such W is the implicit requirement that $W(x, y)$ is supposed to be given as a specified value. For some graph sequences G_n which converge to a finite graph, it is quite straightforward to define the associated W_n as in (1). However, in general, it is quite possible that $W_n(x, y)$ as a function of n approach 0 as n goes to infinity. In such cases, it is better to treat the limit as an operator.

Remark 4. Throughout the paper, $\int F(y)dy$ denotes the usual integration of a function F subject to the Lebesgue measure ν. We will impose the condition that the space of functions that we focus on are Lebesgue measurable and integrable so that all the inner products involving integration make sense. For some other measures, such as μ_n and μ for a graph sequence, as defined in Section 2.1 and 2.2, it can be easily checked that if a function F is Lebesgue measurable and integrable then F is also measurable and integrable subject to μ_n and μ.

2.1 The Laplace operator on a graph

For a weighted graph G_n on n vertices with edge weight $A_n(u, v)$ for vertices u and v , we define the Laplace operator Δ_n to be

$$\Delta_n f(u) = \frac{1}{d_u} \sum_v (f(u) - f(v)) A_n(u, v). \tag{3}$$

for $f : V \to \mathbb{R}$. It is easy to check that

$$\Delta_n = I_n - D_n^{-1} A_n = D_n^{-1/2} \mathcal{L}_n D_n^{1/2}$$

where I_n is the $n \times n$ identity matrix and $\mathcal{L}_n = I_n - D_n^{-1/2} A_n D_n^{-1/2}$ is the symmetric normalized Laplacian (see [13]).

Let μ_n denote the measure defined by $\mu_n(v) = d_v/\mathrm{vol}(G)$ for v in G_n where $\mathrm{vol}(G_n) = \sum_v d_v$. We define an inner product on functions $f, g : V \to \mathbb{R}$ by

$$\langle f, g \rangle_{\mu_n} = \sum_{v \in V} f(v) g(v) \mu_n(v).$$

It is then straightforward to check that

$$\begin{aligned} \sum_{\{u,v\} \in E} \frac{(f(u) - f(v))(g(u) - g(v)) A_n(u, v)}{\mathrm{vol}(G_n)} &= \frac{\sum_u f(u) \sum_{v \sim u} (g(u) - g(v)) A_n(u, v)}{\mathrm{vol}(G_n)} \\ &= \sum_u f(u) (\Delta_n g)(u) \frac{d_u}{\mathrm{vol}(G_n)} \\ &= \langle f, \Delta_n g \rangle_{\mu_n} \end{aligned}$$

and

$$\langle f, \Delta_n g \rangle_{\mu_n} = \langle g, \Delta_n f \rangle_{\mu_n}.$$

If f and g are complex-valued functions, then we have

$$\langle f, \Delta_n g \rangle_{\mu_n} = \overline{\langle g, \Delta_n f \rangle_{\mu_n}}$$

where $\bar{x}$ denotes the complex conjugate of x.

We note that $\langle f, \Delta_n \mathbf{1}\rangle_{\mu_n} = \langle \mathbf{1}, \Delta_n f\rangle_{\mu_n} = 0$, where $\mathbf{1}$ denotes the constant function 1. Therefore, Δ_n has an eigenvalue 0 with an associated eigenfunction $\mathbf{1}$, under the μ_n-norm. The eigenfunctions ϕ_j, for $j = 0, \ldots, n-1$, form an orthogonal basis under the μ_n-norm for G_n. In other words, $D_n^{1/2}\phi_j$ form an orthogonal basis under the usual inner product as eigenvectors for the normalized Laplacian $I_n - D_n^{-1/2} A_n D_n^{-1/2}$. The ϕ_j's are previously called the combinatorial eigenfunctions in [13].

.

2.2 The convergence of degree distributions

Suppose we have a sequence of graphs. For a graph G_n on n vertices, the measure μ_n, defined by $\mu_n(v) = \frac{d_n(v)}{\mathrm{vol}(G_n)}$, is also called the degree distribution of G_n where $d_n(v)$ denotes the degree of v in G_n and $\mathrm{vol}(G_n) = \sum_v d_n(v)$. In general, for a subset X of vertices in G_n, $\mathrm{vol}_{G_n}(X) = \sum_{v \in X} d_n(v)$. In this paper, we focus on graph sequences with convergent degree distributions which we will describe.

For a graph G_n with vertex set V_n consisting of n vertices, we let F_n denote the set of all bijections from V_n to $\{1, 2, \ldots, n\}$.

$$F_n = \{\eta : V_n \to \{1, 2, \ldots, n\}\}. \tag{4}$$

For each $\eta \in F_n$, we let η_n denote the associated partition map $\eta_n : [0,1] \to V_n$, defined by $\eta_n(x) = \eta(u)$ if $x \in ((\eta(u) - 1)/n, \eta(u)/n] = I_{\eta(u)}$. We write $I_{\eta(u)} = I_u$ if there is no confusion. In stead of F_n, it is sometimes convenient to consider

$$\mathcal{F}_n = \{\eta_n : \eta \in F_n\} \subset \{\varphi : [0,1] \to V_n\} \tag{5}$$

Now, for any integrable functions $f, g : [0,1] \to \mathbb{R}$, we define

$$\langle f, g\rangle_{\mu_n, \eta_n} = \int_0^1 f(x) g(x) \mu_n^{(\eta_n)}(x) \tag{6}$$

where μ_n is defined by

$$\int_0^1 F(x) \mu_n^{(\eta_n)}(x) = \int_0^1 F(x) n \mu_n(\eta_n(x)) dx \tag{7}$$

for integrable $F : [0,1] \to \mathbb{R}$. We can then define the associated norm:

$$\|f\|_{\mu_n, \eta_n} \quad = \quad \sqrt{\langle f, f\rangle_{\mu_n, \eta_n}} \tag{8}$$

As a measure on $[0,1]$, $\mu_n^{(\eta_n)}$ satisfies

$$\mu_n(u) = \int_{I_u} \mu_n^{(\eta_n)}(x) \tag{9}$$

and

$$\int_0^1 \mu_n^{(\eta_n)}(x) = 1.$$

For example, for a graph G_5 with degree sequence $(2, 2, 3, 3, 4)$, and suppose the corresponding vertices are denoted by $v_1, \dots, v_5$, then $\mu_n(v_1) = \mu_n(v_2) = 1/7$ and $\mu_n(v_3) = \mu_n(v_4) = 3/14$, etc.

In particular, for a subset $S \subseteq [0, 1]$, we consider the characteristic function $\chi_S(x) = 1$ if $x \in S$ and 0 otherwise. Then for $f = g = \chi_S$, we have

$$\begin{aligned} \langle \chi_S, \chi_S \rangle_{\mu_n, \eta_n} &= \mu_n^{(\eta_n)}(S) \\ &= \int_S \mu_n^{(\eta_n)}(x). \end{aligned} \tag{10}$$

Sometimes we suppress the labeling map η_n and simply write μ_n as the associated measure on $[0, 1]$ if there is no confusion.

For $\epsilon > 0$, we say two graphs G_m and G_n have ϵ-similar degree distributions if

$$\inf_{\theta \in \mathcal{F}_m, \eta \in \mathcal{F}_n} \int_0^1 |\mu_m^{(\theta)}(x) - \mu_n^{(\eta)}(x)| < \epsilon. \tag{11}$$

For a graph sequence $G_n, n = 1, 2, \dots$, we say the degree distribution μ_n is Cauchy, if for any $\epsilon > 0$, there exists $N = N(\epsilon)$ such that for any $m, n \geq N$, the degree distributions of G_m and G_n are ϵ-similar. To see that the degree distributions converge, we use the following arguments:

Lemma 1. *If the degree distribution of the sequence G_n is Cauchy, then there are $\theta_n \in \mathcal{F}_n$ such that the sequence $\mu_n^{(\theta_n)}$ of G_n converges to a limit, denoted by μ. Furthermore μ is unique up to a measure preserving map.*

Proof. For each positive integer j, we set $\epsilon_j = 2^{-j}$, and let $N(\epsilon_j)$ denote the least integer such that for $m, n \geq N(\epsilon_j)$, G_m and G_n have ϵ_j-similar degree distributions. To simplify the notation, we write $M(j) = N(\epsilon_j)$.

We first choose an arbitrary permutation $\eta_{M(1)}$ and then by induction define permutations $\eta_{M(j)}$'s, for $j > 1$ using (11) so that

$$\int_0^1 |\mu_{M(j)}^{(\theta_{M(j)})}(x) - \mu_{M(j+1)}^{(\theta_{M(j+1)})}(x)| < \epsilon_j.$$

For each $n \in [M(j), M(j+1))$, we choose the permutation η_n such that

$$\int_0^1 |\mu_n^{(\theta_n)}(x) - \mu_{M(j)}^{(\theta_{M(j)})}(x)| < \epsilon_j.$$

Claim: the sequence of $\mu_n^{(\theta_n)}$, for $n = 1, 2, \ldots$ is Cauchy.
To prove the claim, we see that for any $m, n \geq M(j)$ satisfying $n \in [M(j), M(j+1))$ and $m \in [M(k), M(k+1))$ with $j \leq k$, we have

$$\begin{aligned}
& \int_0^1 \mid \mu_n^{(\theta_n)}(x) - \mu_m^{(\theta_m)}(x) \mid \\
\leq & \int_0^1 \mid \mu_n^{(\theta_n)}(x) - \mu_{M(j)}^{(\theta_{M(j)})}(x) \mid + \int_0^1 \mid \mu_{M(j)}^{(\theta_{M(j)})}(x) - \mu_{M(j+1)}^{(\theta_{M(j+1)})}(x) \mid + \ldots \\
& + \int_0^1 \mid \mu_{M(k-1)}^{(\theta_{M(k-1)})}(x) - \mu_{M(k)}^{(\theta_{M(k)})}(x) \mid + \int_0^1 \mid \mu_{M(k)}^{(\theta_{M(k+1)})}(x) - \mu_m^{(\theta_m)}(x) \mid \\
\leq & \ 2\epsilon_j + \epsilon_{j+1} + \ldots + \epsilon_{k-1} + 2\epsilon_k \\
= & \ 3\epsilon_j
\end{aligned}$$

and the Claim is proved.

To show that the sequence $\mu_n^{(\theta_n)}$ converges, we define $\mu(S)$ for any measurable subset $S \subseteq [0, 1]$ as follows:

$$\begin{aligned}
\mu(S) &= \lim_{n \to \infty} \mu_n^{(\theta_n)}(S) \\
&= \lim_{n \to \infty} \int_S \mu_n^{(\theta_n)}(x).
\end{aligned}$$

Since $\mu_n^{(\theta_n)}$ is Cauchy, the above limit exists and $\mu(S)$ is well defined. Furthermore, for any measure preserving map τ, $\mu \circ \tau$ is the limit of $\mu_n^{(\theta_n \circ \tau)}$. Thus, μ is unique up to a measure preserving map.

To see that μ is a probabilistic measure, we note that for any $\epsilon > 0$, there is some n such that

$$\begin{aligned}
\mid \int_0^1 \mu(x) - 1 \mid &= \mid \int_0^1 \mu(x) - \int_0^1 \mu_n^{(\eta_n)}(x) \mid \\
&= \int_0^1 \mid \mu(x) - \mu_n^{(\eta_n)} \mid \\
&\leq \epsilon.
\end{aligned}$$

Lemma 1 is proved. □

Remark 5. Since we are dealing with exchangeable operators, the measures μ can be regarded as the equivalence class of probabilistic measures where two measures φ, φ' are said to be equivalent if there is a Lebesgue measure preserving bijection τ on $[0, 1]$ such that $\varphi = \varphi' \circ \tau$.

Remark 6. An alternative proof for the convergence μ_n is due to Stephen Young [45] which is simpler but the resulted limit μ is not necessarily exchangeable. For each n, suppose we choose η_n such that $\mu_n^{(\eta_n)}$ is a non-decreasing function on $[0,1]$. By using the fact that for $x_1 < x_2$ and $y_1 < y_2$, we have $|x_1 - y_1| + |x_2 - y_2| \le |x_1 - y_2| + |x_2 - y_1|$, it follows that

$$\inf_{\theta\in\mathcal{F}_m,\eta\in\mathcal{F}_n} \int_0^1 |\mu_m^{(\theta)}(x) - \mu_n^{(\eta)}(x)| = \int_0^1 |\mu_m^{(\eta_m)}(x) - \mu_n^{(\eta_n)}(x)|$$

Thus the sequence $\mu_n^{(\eta_n)}$ is Cauchy and therefore converges to a limit μ.

We note that two different graphs G and H both on n vertices can have the same degree distribution measure μ_n but G and H have different degree sequences. For example, G is a k-regular graph and H is a k'-regular graph where $k \neq k'$. In this case, $\mu_n(v) = 1/n$ for any vertex v and $\mu_n(x) = 1$ for any $x \in [0,1]$. To define the convergence of graph sequences, we need to take into account the volume $\text{vol}(G) = \sum_v d_v$ of G.

2.3 The spectral distance

Suppose we consider two graphs G_m and G_n on m and n vertices, respectively. Their associated Laplace operators are denoted by Δ_m and Δ_n, respectively. If $m \neq n$, Δ_m and Δ_n have different sizes. In order to compare two given matrices, we need some definitions.

In $G_n = (V_n, E_n)$, for $\eta_n \in \mathcal{F}_n$ (as described in (4)), the operator $\Delta_n^{(\eta)}$ is acting on an integrable function $f : [0,1] \to \mathbb{R}$ by

$$\begin{aligned} \Delta_n^{(\eta)} f(x) &= \frac{n}{d_n(\eta_n(x))} \int_0^1 \big(f(x) - f(y)\big) W_n^{(\eta_n)}(x,y)dy \\ &= f(x) - \frac{n}{d_n(\eta_n(x))} \int_0^1 W_n^{(\eta_n)}(x,y) f(y) dy \end{aligned} \tag{12}$$

where $W_n^{(\eta_n)} \in \mathcal{W} = [0,1] \times [0,1]$ is associated with the adjacency matrix A_n by $W^{(\eta_n)}(x,y) = A_n(\eta_n(x), \eta_n(y))$. Here we require f to be Lebesgue measurable and therefore is also μ_n-measurable. In the remainder of the paper, we deal with functions that are Lebesgue integrable on $[0,1]$. We note that for any two permutations $\theta, \eta \in F_n$, $W^{(\theta_n)}$ is equivalent to $W^{(\eta_n)}$ as exchangeable operators in $\mathcal{W}^*$, defined in (3).

For a Lebesque integrable function $f : [0,1] \to \mathbb{R}$, we consider

$$\begin{aligned} \langle f, \Delta_n^{(\eta_n)} g\rangle_{\mu_n,\eta_n} &= \int_0^1 f(x)\big(\Delta_n^{(\eta_n)} g(x)\big)\mu_n^{(\eta_n)}(x) \\ &= \int_0^1 \int_0^1 \frac{n}{d_n(\eta_n(x))} f(x)\big(g(x) - g(y)\big) W_n^{(\eta_n)}(x,y)\ dy\ \mu_n^{(\eta_n)}(x) \end{aligned}$$

Using (7), we have

$$\begin{aligned}\langle f, \Delta_n^{(\eta_n)} g\rangle_{\mu_n,\eta_n} &= \frac{n^2}{\mathrm{vol}(G_n)} \int_0^1 \int_0^1 f(x)\big(g(x)-g(y)\big) W_n^{(\eta_n)}(x,y)dxdy \\ &= \frac{n^2}{2\mathrm{vol}(G_n)} \int_0^1 \int_0^1 \big(f(x)-f(y)\big)\big(g(x)-g(y)\big) W_n^{(\eta_n)}(x,y)dxdy.\end{aligned}$$

In particular,

$$\langle f, \Delta_n^{(\eta_n)} f\rangle_{\mu_n,\eta_n} = \frac{n^2}{2\mathrm{vol}(G_n)} \int_0^1 \int_0^1 \big(f(x)-f(y)\big)^2 W_n^{(\eta_n)}(x,y)dxdy. \quad (13)$$

Remark 7. The above inner products are invariant subject to any choice of measure preserving maps τ. Namely, if we define $f\circ\tau(x) = f(\tau(x))$, then

$$\langle f, \Delta_n^{(\eta_n)} f\rangle_{\mu_n,\eta_n} = \langle f\circ\tau, \Delta_n^{(\eta_n\circ\tau)} f\rangle_{\mu_n\circ\tau,\eta_n\circ\tau}. \quad (14)$$

For an operator M acting on the space of integrable functions $f:[0,1]\to\mathbb{R}$, we say M is *exchangeable* if for any measure preserving map τ, we have

$$\langle f, Mg\rangle = \langle f\circ\tau, M_\tau(g\circ\tau)\rangle$$

where M_τ is defined by $M_\tau h(x,y) = M(h(\tau^{-1}(x)), h(\tau^{-1}(y)))$. Clearly, Δ_n is an exchangeable operator.

For an integrable function $f:[0,1]\to\mathbb{R}$ and $\eta_n\in\mathcal{F}_n$, we define $\tilde{f}_n:[0,1]\to\mathbb{R}$, for $x\in I_u$, as follows:

$$\begin{aligned}\tilde{f}_n(x) = \frac{\int_{I_u} f(y)\mu_n^{(\eta_n)}(y)dy}{\int_{I_u}\mu_n^{(\eta_n)}(y)dy} &= \frac{\int_{I_u} f(y)\mu_n^{(\eta_n)}(y)dy}{\mu_n(u)} \\ &= \frac{\int_0^1 I_n^{(\eta_n)}(x,y) f(y)\mu_n^{(\eta_n)}(y)dy}{\mu_n(u)} \quad (15)\end{aligned}$$

where I_n is the $n\times n$ identity matrix as defined in Section 2.1. Note that $\tilde{f}_n(x) = \tilde{f}_n(z)$ if $\eta_n(x) = \eta_n(z)$. For u in $V(G_n)$, we write $\tilde{f}_n(u) = \tilde{f}_n(x)$ where $x\in I_u$.

f and $\tilde{f}$ are related as follows:

Lemma 2.
(i) For $x\in I_u$, and $\eta_n\in\mathcal{F}_n$,

$$\begin{aligned}\Delta_n^{(\eta_n)} f(x) &= \frac{n}{d_n(\eta_n(x))}\sum_v \int_{y\in I_v} (f(x)-f(y))A_n(\eta_n(x),v)dy \\ &= \frac{1}{d_n(u)}\sum_v (f(u)-\tilde{f}_n(v))A_n(u,v) \\ &= \Delta_n\tilde{f}_n(u) + (f(x)-\tilde{f}_n(u)).\end{aligned}$$

(ii) For $f, g : [0,1] \to \mathbb{R}$,

$$\langle f, \Delta_n^{(\eta_n)} f\rangle_{\mu_n,\eta_n} \quad = \quad \langle \tilde{f}_n, \Delta_n \tilde{f}_n\rangle_{\mu_n} + \|f - \tilde{f}_n\|^2_{\mu_n,\eta_n}. \tag{16}$$

The proof of (i) follows from (3) and (15). (ii) follows from (i) and (13) by straightforward manipulation.

Remark 8. In this paper, we define inner products and norms on the space of integrable functions defined on $[0,1]$, as seen in (6) and (8). Consequently, the last term in (16) approaches 0 as n goes to infinity. Namely,

$$\|f - \tilde{f}_n\|^2_{\mu_n,\eta_n} \to 0 \quad \text{as } n \to \infty$$

if f is integrable. This implies that the graph Laplacian Δ_n for G_n acting on the space of functions defined on V_n can be approximated by $\Delta_n^{(\eta_n)}$ acting on the space of functions defined on $[0,1]$ with the exception for the function f with $\|\Delta_n f\|_{\mu_n,\eta_n}$ is too close to 0, while f is orthogonal to the eigenfunction associated with eigenvalue 0. The case of a path P_n is one such example and in fact, the graph sequence of paths P_n does not converge under the spectral distance that we shall define. In order to make sure that $\Delta_n^{(\eta_n)}$ closely approximates Δ_n, there are two ways to proceed. We can restrict (implicitly) ourselves to graph sequences G_n with the least nontrivial eigenvalue λ_1 of Δ_n greater than some absolute positive constant (as done in this paper). An alternative way is to consider general labeling space Ω_0 other than $[0,1]$ and impose further conditions on the space of functions defined on Ω_0 (which will be treated in a subsequent paper).

For a graph sequence $G_n = (V_n, E_n)$, where $n = 1, 2, \ldots$, we say the sequence of the Laplace operators Δ_n is Cauchy if for any $\epsilon > 0$ there exists N such that for $m, n \geq N$, there exist $\theta_m \in \mathcal{F}_m$, $\theta_n \in \mathcal{F}_n$ such that the following holds:
(i) The associated measures $\mu_m^{(\theta_m)}$ and $\mu_m^{(\theta_m)}$ satisfy

$$\int_0^1 |\mu_m^{(\theta_m)}(x) - \mu_n^{(\theta_n)}(x)| < \epsilon.$$

(ii) The Laplace operators associated with G_m and G_n satisfy

$$\left| \frac{\langle f, \Delta_m^{(\theta_m)} g\rangle_{\mu_m,\theta_m}}{\|f\|_{\mu_m,\theta_m}\|g\|_{\mu_m,\theta_m}} - \frac{\langle f, \Delta_n^{(\theta_n)} g\rangle_{\mu_n,\theta_n}}{\|f\|_{\mu_n,\theta_n}\|g\|_{\mu_n,\theta_n}} \right| < \epsilon$$

for integrable f, g defined on $[0,1]$ and we write

$$d(\Delta_m, \Delta_n)_{\mu_m,\mu_n} < \epsilon \tag{17}$$

where μ_m, μ_n denote the degree distributions of G_m, G_n, respectively.

Remark 9. *We note that the spectral distance here is invariant subject to any choices of measure preserving maps. In fact, for any measure preserving map* τ*, it follows from the definition and that* $d(\Delta_m, \Delta_n)_{\mu_m,\mu_n} < \epsilon$ *if and only if* $d(\Delta_m, \Delta_n)_{\mu_m\circ\tau,\mu_n\circ\tau} < \epsilon$.

Suppose the sequence of graphs $G_n = (V_n, E_n)$ have degree distributions μ_n converging to μ as above. Then (17) can be simplified. The inequality in (17) can be replaced by an equivalent condition

$$d(\Delta_m, \Delta_n)_\mu < \epsilon$$

which can be described by there exists N such that for $m, n \geq N$, there exist $\theta_m \in F_m$, $\theta_n \in F_n$ such that the Laplace operators associated with G_m and G_n satisfy

$$\left|\langle f, (\Delta_m^{(\theta_m)} - \Delta_n^{(\eta_n)})g\rangle_\mu\right| < \epsilon \tag{18}$$

for integrable $f, g : [0, 1] \to \mathbb{R}$ with $\|f\|_\mu = \|g\|_\mu = 1$.

For an operator M on $[0, 1]$ we can define spectral μ-norm, defined by

$$\|M\|_\mu^2 = \sup_{f,g} |\ \langle f, Mg\rangle_\mu\ |$$

where $f, g : [0, 1] \to \mathbb{R}$ range over integrable functions satisfy $\|f\|_\mu = \|g\|_\mu = 1$. We are ready to examine the convergence of a graph sequence under the spectral distance.

Theorem 1. *For a graph sequence* $G_n = (V_n, E_n)$*, where* $n = 1, 2, \ldots$*, suppose the sequence of the Laplace operators* Δ_n *is Cauchy, then for each* n*, there are permutations* $\theta_n \in F_n$ *such that the sequence of* $\Delta_n^{(\theta_n)}$ *converges to an exchangeable operator* Δ *and the measure* $\mu_n^{(\theta_n)}$ *of* G_n*'s converge to* μ *where* Δ *satisfies*

$$\int_0^1 f(x)\Delta g(x)\mu(x) = \lim_{n\to\infty} \langle f, \Delta_n^{(\theta_n)} g\rangle_{\mu_n} \tag{19}$$

for any two integrable functions $f, g : [0, 1] \to \mathbb{R}$.

Proof. For each positive integer j, we set $\epsilon_j = 2^{-j}$, and let $N(\epsilon_j)$ denote the least integer such that for $m, n \geq N(\epsilon_j)$, (17) holds for ϵ_j. To simplify the notation, we write $M(j) = N(\epsilon_j)$.

We first choose an arbitrary permutation $\eta_{(1)} \in F_{M(j)}$ and then by induction define permutations $\theta_{(j)} \in F_{M(j)}$'s, for $j > 1$, using (17) so that

$$d(\Delta_{M(j)}, \Delta_{M(j+1)})_{\mu_{M(j)}^{(\theta_{(j)})},\mu_{M(j+1)}^{(\theta_{(j+1)})}} < \epsilon_j.$$

We can assume the associated measure for $\theta_{(j)}$ is non-decreasing since we can simply adjust by choosing measure preserving maps.

For each $n \in [M(j), M(j+1))$, we choose the permutation θ_n such that

$$d(\Delta_n, \Delta_{M(j)})_{\mu_n^{(\theta_n)}, \mu_{M(j)}^{(\theta_{(j)})}} < \epsilon_j.$$

We will use a similar method as in Lemma 1 to prove the following:

Claim 1: The sequence of $\Delta_n^{(\theta_n)}$, for $n = 1, 2, \ldots$ is Cauchy.

To prove the claim, we see that for any $m, n \geq M(j)$ satisfying $n \in [M(j), M(j+1))$ and $m \in [M(k), M(k+1))$ with $j \leq k$, we have

$$\begin{aligned}
& d(\Delta_m, \Delta_n)_{\mu_n^{(\theta_n)}, \mu_m^{(\theta_m)}} \\
\leq\ & d(\Delta_n, \Delta_{M(j)})_{\mu_n^{(\theta_n)}, \mu_{M(j)}^{(\theta_{(j)})}} + \ldots \\
& + d(\Delta_{M(k-1)}, \Delta_{M(k)})_{\mu_{M(k-1)}^{(\theta_{(k-1)})}, \mu_{M(k)}^{(\theta_{(k)})}} + d(\Delta_{M(k)}, \Delta_m)_{\mu_{M(k)}^{(\theta_{(k)})}, \mu_m^{(\theta_n)}} \\
\leq\ & 2\epsilon_j + \epsilon_{j+1} + \ldots + \epsilon_{k-1} + 2\epsilon_k \\
=\ & 3\epsilon_j
\end{aligned}$$

and Claim 1 is proved.

Claim 2: The sequence of $\mu_n^{(\theta_n)}$ is Cauchy and therefore converges to a limit μ. To prove Claim 2, we will first show that for any $\epsilon > 0$, $m, n \geq N(\epsilon_j)$, and any subset $S \subset [0, 1]$, we have $|\mu_m^{(\theta_m)}(S) - \mu_n^{(\theta_n)}(S)| \leq 6\epsilon_j$.

From the proof of Claim 1, we know that $d(\Delta_m^{(\theta_m)}, \Delta_n^{(\theta_n)}) \leq 3\epsilon_j$, which implies, by choosing $f = \chi_S$ and $g = \mathbf{1}$ in (47) and (**??**),

$$\begin{aligned}
3\epsilon_j &\geq d(\Delta_m^{(\theta_m)}, \Delta_n^{(\theta_n)}) \\
&\geq \left| \sqrt{\mu_m^{(\theta_m)}(S)} - \sqrt{\mu_n^{(\theta_n)}(S)} \right| \\
&\geq \frac{\left| \mu_m^{(\theta_m)}(S) - \mu_n^{(\theta_n)}(S) \right|}{\sqrt{\mu_m^{(\theta_m)}(S)} + \sqrt{\mu_n^{(\theta_n)}(S)}} \\
&\geq \frac{1}{2} \left| \mu_m^{(\theta_m)}(S) - \mu_n^{(\theta_n)}(S) \right|.
\end{aligned}$$

To show that $\mu_n^{(\theta_n)}$ is Cauchy, we set $S = \{x : \mu_n^{(\theta_n)}(x) > \mu_m^{(\theta_m)}(x)\}$. Then,

$$\begin{aligned}
\int_0^1 | \mu_n^{(\theta_n)}(x) - \mu_m^{(\theta_m)}(x) | &= 2 \int_S | \mu_n^{(\theta_n)}(x) - \mu_m^{(\theta_m)}(x) | + \int_{\bar{S}} | \mu_n^{(\theta_n)}(x) - \mu_m^{(\theta_m)}(x) | \\
&= 2 \, | \mu_n^{(\theta_n)}(S) - \mu_m^{(\theta_m)}(S) | \\
&\leq 12\epsilon_j.
\end{aligned}$$

Claim 2 is proved.

Now, we can define the operator Δ:

$$\begin{aligned} \langle f, \Delta g \rangle &= \int_0^1 f(x)\Delta g(x)\mu(x) && (20)\\ &= \lim_{n\to\infty} \langle f, \Delta_n^{(\theta_n)} g\rangle_{\mu_n} && (21)\end{aligned}$$

for any two integrable functions $f, g : [0,1] \to \mathbb{R}$.

Combining Claims 1 and 2, the sequence $\Delta_n^{(\theta_n)}$ converges to a limit Δ. $\square$

For a graph sequence G_n, where $n = 1, 2, \ldots$, the Laplace operator Δ_n of G_n and $W_{G_n} \in \mathcal{W}^*$ are related as follows: For functions $f, g : [0,1] \to \mathbb{R}$, by using (7) we have

$$\begin{aligned} \langle f, (I - \Delta_n^{(\eta_n)})g\rangle_{\mu_n,\eta} &= \int_0^1 f(x)\big((I-\Delta_n^{(\eta_n)})g(x)\big)\mu_n^{(\eta_n)}(x)\\ &= \int_0^1\int_0^1 \frac{n}{d_n(\eta_n(x))} f(x)g(y)W_n^{(\eta_n)}(x,y)dy\mu_n^{(\eta_n)}(x)\\ &= \frac{n^2}{\mathrm{vol}(G_n)}\int_0^1\int_0^1 f(x)g(y)W_n^{(\eta_n)}(x,y)dxdy\end{aligned}$$

although the existence of the limit of W_{G_n} is not necessarily required.

There are similarities between Δ and the previous definitions for graph limits (as defined in [38]) but the scaling is different as seen below:

$$\begin{aligned} \int_0^1 f(x)\big((I-\Delta)g(x)\big)\mu(x) &= \lim_{n\to\infty}\int_0^1 f(x)\big((I-\Delta_n^{(\eta_n)})g\big)(x)\mu_n^{(\eta_n)}(x)\\ &= \lim_{n\to\infty}\langle f, \frac{n^2}{\mathrm{vol}(G_n)} W_n g\rangle. && (22)\end{aligned}$$

Suppose the graph sequence have volume $\mathrm{vol}(G_n)$ converging to a function Φ. Then we have

$$\int_0^1 f(x)\big((I-\Delta)g(x)\big)\mu(x) = \lim_{n\to\infty}\langle f, \frac{n^2}{\mathrm{vol}(G_n)} W_{G_n} g\rangle \qquad (23)$$

Thus, the Laplace operator Δ as a limit of Δ_n is essentially the identity operator minus a scaled multiple of the limit W. We state here the following useful fact which follows from Theorem 1:

Lemma 3. *For a sequence of graphs G_n, for $n = 1, 2, \ldots$, with degree distributions μ_n converging to μ, the associated Laplace operators Δ_n converges to Δ satisfying*

$$\langle \chi_S, (I-\Delta)\mathbf{1}\rangle_\mu = \mu(S) \geq 0, \qquad (24)$$

$$\textit{and} \quad \langle \chi_S, (I-\Delta)\chi_T\rangle_\mu \geq 0 \qquad (25)$$

for any integrable subsets $S, T \subseteq [0,1]$ where $\mathbf{1}$ is the constant function assuming the value 1.

Proof. The proof of (24) follows from the fact that

$$\begin{aligned} \langle \chi_S, (I-\Delta)\mathbf{1}\rangle_\mu &= \lim_{n\to\infty} \langle \chi_S^{(n)}, (I-\Delta_n)\mathbf{1}\rangle_{\mu_n} \\ &= \lim_{n\to\infty} \langle \chi_S^{(n)}, \mathbf{1}\rangle_{\mu_n} \\ &= \lim_{n\to\infty} \mu_n(S^{(n)}) \\ &= \mu(S). \end{aligned}$$

To see (25), we note that for any two vertices u, v in G_n, $\langle \chi_u, (I-\Delta_n)\chi_v\rangle_\mu = A_n(u,v)/\mathrm{vol}(G_n) \geq 0$. □

2.4 Defining the graphlets

Using the convergence definitions in the previous subsections, we define graphlets as the limit of a graph sequence

$$G_1, G_2, \ldots, G_n, \cdots \to \mathcal{G}(\Omega, \Delta) \tag{26}$$

which satisfies the following conditions:

1. The degree distributions of G_n introduce measures μ_n on Ω and μ_n converges to a measure μ for Ω as in (11).

2. The discrete Laplace operators Δ_n for G_n converges to Δ as an operator on Ω under the spectral distance using the μ-norm as in (9) and (19).

3. The volume $\mathrm{vol}(G_n)$ of G_n is increasing in n.

Several examples of graphlets will be given in the next section.

Remark 10. One advantage of the graphlets $\mathcal{G}(\Omega, \Delta)$ is the fact that the eigenvectors of graphs in the graph sequences can be approximated by eigenvectors of Δ. In other words, eigenvectors of Δ can be used as universal basis for all graphs in graph sequences in the graphlets $\mathcal{G}(\Omega, \Delta)$.

Remark 11. In the other direction, graphs in graph sequences in graphlets $\mathcal{G}(\Omega, \Delta)$ can be viewed as a scaling for discretization of Ω and Δ. If two different graph sequences converge to the same graphlets, they can be viewed as giving different scaling for discretization.

Remark 12. Another way to describe a graphlets is to view $\mathcal{G}(\Omega, \Delta)$ as the limit of graphlets $\mathcal{G}(\Omega_n, \Delta_n)$. Here Ω_n can be described as a measure space under a measure μ_n as follows. The elements in Ω_n, (the same as that of Ω, labelled by $[0,1]$) is the union of n parts, denoted by I_v, indexed by vertices v of G_n. The degree of v satisfies

$$d_n(v) \quad \approx \quad \mathrm{vol}(G_n) \int_{I_v} \mu(x).$$

The Laplace operator Δ_n can be defined by using the adjacency entry $A_n(u,v) = W_n(x,y)$ for $x \in I_u$ and $y \in I_v$. Namely, $\Delta_n(x,y) = I_n(x,y) - W_n(x,y)/d_x$. The graphlets $\mathcal{G}(\Omega, \Delta)$ as the limit of $\mathcal{G}(\Omega_n, \Delta_n)$ specifies the incidence quantitiy between any two integrable subsets S and T in Ω. For an integrable $S \subseteq \Omega$, we let χ_S denote the characteristic function of S, which assume the value 1 on S, and 0 otherwise. In Ω_n, the incidence quantity between S and T, denoted by $\mathcal{E}_n(S,T)$ satisfies:

$$\mathcal{E}_n(S,T) = \mathrm{vol}(G_n) \int_0^1 \chi_S(x)\big((I - \Delta_n)\chi_T(x)\big)\mu_n(x). \tag{27}$$

In particular, for $S = T$,

$$E_n(S,S) \quad \approx \quad \mathrm{vol}(G_n)\big(\mu(S) - \mu(\partial(S))\big)$$

where the boundary $\partial(S)$ of S satisfies

$$E_n(S,\bar{S}) \approx \mathrm{vol}(G_n)\mu(\partial(S)) = \mathrm{vol}(G_n) \int_0^1 \chi_S(x)\Delta\chi_S(x)\mu(x).$$

3 Examples of graphlets

We here consider several examples of graphlets $\mathcal{G}(\Omega, \Delta)$ which are formed from graph sequences G_n, for $n = 1, 2, \ldots$. We will illustrate that the eigenfunctions of Δ can be used to serve as a universal basis for graphs G_n. The discretized adaptation of graphlets will be called "*lifted* graphlets" for G_n, which are good approximations for the actual eigenfunctions in G_n as n approaches infinity. In some cases, the lifted graphlets using Δ are fewer than the number of eigenfunctions in G_n and in other cases, there are more eigenfunctions of Δ than those of G_n. We will describe a universal basis for G_n, as the union of two parts, including the *primary* series (which are the lifted graphlets) and *complementary* series (which are orthogonal to the primary series). In a way, we will see that the primary series captures the main structures of the graphs while the complementary series reflect the "noise" toward the convergence. Before we proceed, some clarifications are in order.

- The notion of orthogonality refers to the usual inner product unless we specify other modified inner products such as the μ-product $\langle \cdot, \cdot \rangle_\mu$ or the μ_n-product. Sometimes, it is more elegant to use eigenfunctions that are orthogonal under the μ-norm. However, when we are dealing with a finite graph G_n in a graph sequence, we sometimes wish to use only what we know about the finite graph G_n and perhaps the existence of the limit without the knowledge of the behavior of the limit (such as μ). In such cases, we will use the usual inner product.

- The universal bases are for approximating the eigenfunctions of the normalized Laplacian of G_n. In a graph G_n, its Laplace operator $\Delta_n = I - D_n^{-1}A_n$ is not symmetric in general since the left and right eigenfunctions are not necessarily the same. The universal basis is used for approximating the eigenfunctions of the normalized Lapalcian

$$\mathcal{L}_n = I - D_n^{-1/2} A_n D_n^{-1/2},$$

which is equivalent to Δ_n and is symmetric. Thus, $\mathcal{L}$ has orthogonal eigenfunctions.

3.1 Dense graphlets

Suppose we have a sequence of dense graphs G_n, for $n = 1, 2, \ldots$, with $\mathrm{vol}(G_n) = 2|E(G_n)| = c_n n^2$ where the c_n converge to a constant $c > 0$. In this case, the μ-norm is equivalent to other norms such as the cut-norm and subgraph-norm in [9]. By using the regularity lemma, the graphlets $\mathcal{G}(\Omega, \Delta)$ of a dense graph sequence is of a finite type. In other words, there is a graph H on h vertices where h is a constant (independent of n) such that $\Omega = \Omega_H$ is taken to be a partition of $[0, 1]$ into h intervals of the same length. Let $\varphi_1, \ldots, \varphi_h$ denote the eigenfunctions of H.

For $n = hm$ and $m \in \mathbb{Z}$, we will describe a basis for a graph G_n. The primary eigenfunctions can be written as

$$\phi_j^{(n)}(v) \quad = \quad \varphi_j(\lceil v/m \rceil) \quad \text{where } v \in \{1, 2, \ldots, n\} \text{ and } j = 1, \ldots, h,$$

while the complementary eigenfunctions consist of $n - h = (m-1)h$ eigenfunctions as follows: For $1 \le a \le h, 1 \le b \le m - 1$,

$$\phi_{a,b}^{(n)}(a'm + b') \quad = \quad \begin{cases} e^{2\pi i b b'/m} & \text{if } a' + 1 = a, \\ 0 & \text{otherwise.} \end{cases}$$

3.2 Quasi-random graphlets

Originally, quasi-randomness is an equivalent class of graph properties that are shared by random graphs (see [19]). In the language of graph limits, quasi-random graph properties with edge density 1/2 can be described as a graph sequence G_n, for $n = 1, 2, \ldots$, converging to graphlets $\mathcal{G}(\Omega, \Delta)$ where $\Omega = [0, 1]$ and $\Delta(x, y) = 1/2$, for $x \neq y$ and $\mu(x) = \mu(y)$ for all x, y. Compared with the original equivalent quasi-random properties for G_n (included in parentheses), the quasi-random graphlets with edge density 1/2 satisfies the following equivalent statements for the graph sequence G_n where $n = 1, 2, \ldots$.

(1) The graph sequence G_n converges to graphlets $\mathcal{G}(\Omega, \Delta)$ in the spectral distance.
(**The eigenvalue property**: *The adjacency matrix of G_n on n vertices has one eigenvalue $n/2 + o(n)$ with all other eigenvalues $o(n)$.)*

(2) The graph sequence G_n converges to graphlets $\mathcal{G}(\Omega, \Delta)$ in the cut-distance.
(**The discrepancy property**: *For any two subsets S and T of the vertex set of G_n, there are $|S|\cdot|T|/2+o(n^2)$ ordered pairs (u, v) with $u \in S, v \in T$ and $\{u, v\}$ being an edge of G_n.)*

(3) The graph sequence G_n converges to graphlets $\mathcal{G}(\Omega, \Delta)$ in the C_4-count-distance.
(**The co-degree property**: *For all but $o(n^2)$ pairs of vertices u and v in G_n, u and v have $n/4 + o(n)$ common neighbors.)*
*(***The trace property***: The trace of the adjacency matrix to the 4th power is $n^4/16 + o(n^4)$.)*

(4) The graph sequence G_n converges to graphlets $\mathcal{G}(\Omega, \Delta)$ in the subgraph-count-distance.
(**The subgraph-property**: *For fixed $k \geq 4$ and for any H on k vertices and l edges, the number of occurrence of H as subgraphs in G_n is $n^k/2^l + o(n^k)$.)*

(5) The graph sequence G_n converges to graphlets $\mathcal{G}(\Omega, \Delta)$ in the homomorphism-distance.
(**The induced-subgraph-property**: *For fixed $k \geq 4$ and for any H on k vertices, the number of occurrence of H as induced subgraphs in G_n is $n^k/2^{\binom{k}{2}} + o(n^k)$.)*

For a quasi-random graph sequence, the primary graphlets for G_n consists of the all 1's vector $\mathbf{1}$ and the complementary ones are irrelevant in the sense that they can be any arbitrarily chosen orthogonal functions since all eigenvalues except for one approach zero. In other words, Δ as the limit of G_n only has one nontrivial eigenfunction.

The generalization of quasi-randomness to sparse graphs and to graphs with general degree distributions [17, 18] can also be described in the framework of graphlets. In the previous work on quasi-random graphs with given degree distributions, the results are not as strong since additional conditions are required in order to overcome various difficulties [17, 18]. By using graph limits, such obstacles and additional conditions can be removed. In Section 5, we will give a complete characterization for quasi-random graphlets with any given general degree distribution which include the sparse cases.

3.3 Bipartite quasi-random graphlets

A bipartite quasi-random graphlets $\mathcal{G}(\Omega, \Delta)$ can be described as follows: Ω is be partitioned into two parts A and B while $W(x, y)$ is equal to some constant ρ if $(x \in A, y \in B)$ or $(x \in B, y \in A)$, and 0 otherwise. There are two nontrivial eigenvalues of $I - \Delta$, namely, 1 and -1. The eigenfuction ϕ_0 associated with eigenvalue 1 assumes the value $\phi_0(x) = 1/\sqrt{\mu(A)}$ for $x \in A$ and $\phi_0(y) = 1/\sqrt{\mu(B)}$ for $y \in B$. The eigenfunction ϕ_1 associated with eigenvalue -1 is defined by $\phi_1(x) = 1/\sqrt{\mu(A)}$ for $x \in A$ and $\phi_1(y) = -1/\sqrt{\mu(B)}$ for $y \in B$.

The bipartite version of quasi-random graphs is useful in the proof of the regularity lemma [44]. Bipartite quasi-random graphlets, as well as quasi-random graphlets, serve as the basic building blocks for general types of graphlets. More on this will be given in Sections 7 and 8.

3.4 Graphlets of bounded rank

A quasi-random sequence is a graph sequence which converges to a graphlets of rank 1 as we will see in this section. We will further consider the generalization of graph sequences which converge to a graphlets of rank k. This will be further examined in Sections 7 and 8.

4 The spectral distance and the discrepancy distance

4.1 The cut distance and the discrepancy distance

In previous studies of graph limits, a so-called cut metric that is often used for which the distance of two graphs G and H which share the same set of vertices V is measured by the following (see [9, 30]).

$$\text{cut}(G, H) = \frac{1}{|V|^2} \sup_{S,T \subseteq V} |E_G(S, T) - E_H(S, T)| \tag{28}$$

where $E_G(S, T)$ denotes the number of ordered pairs (u, v) where u is in S, v is in T and $\{u, v\}$ is an edge in G.

We will define a discrepancy distance which is similar to but different from the above cut distance. For two graphs G and H on the same vertex set V, the discrepancy distance, denoted by $\text{disc}(G, H)$ is defined as follows:

$$\text{disc}(G, H) = \sup_{S,T \subseteq V} \left| \frac{E_G(S, T)}{\sqrt{\text{vol}_G(S)\text{vol}_G(T)}} - \frac{E_H(S, T)}{\sqrt{\text{vol}_H(S)\text{vol}_H(T)}} \right|. \tag{29}$$

We remark that the only difference between the cut distance and the discrepancy distance is in the normalizing factor which will be useful in the proof later.

For two graphs G_m and G_n with m and n vertices respectively, we use the labeling maps θ_n and η_n to map $[0,1]$ to the vertices of G_m and G_n, respectively. We define the measures μ_m and μ_n on $[0,1]$ using the degree sequences of G_m and G_n repectively, as in Section 2.2. From the definitions and substitutions, we can write:

$$E_{G_n}(S,T) = \mathrm{vol}(G_n)\langle \chi_S, (I-\Delta_n)\chi_T\rangle_{\mu_n,\theta_n}. \tag{30}$$

Therefore the discrepancy distance in (29) can be written in the following general format:

$$\begin{aligned}&\mathrm{disc}(G_m,G_n)\\ &= \inf_{\theta_m\in\mathcal{F}_m,\eta_n\in\mathcal{F}_n}\ \sup_{S,T\subseteq[0,1]}\left|\frac{\langle \chi_S,(I-\Delta_m)\chi_T\rangle_{\mu_m,\theta_n}}{\sqrt{\mu_m(S)\mu_m(T)}} - \frac{\langle \chi_S,(I-\Delta_n)\chi_T\rangle_{\mu_n,\eta_n}}{\sqrt{\mu_n(S)\mu_n(T)}}\right|\end{aligned} \tag{31}$$

where S,T range over all integrable subsets of $[0,1]$. We can rewrite (30) as follows.

$$E_{G_n}(S,T) \quad = \quad \mathrm{vol}(G_n)\int_{x\in\Omega}\chi_S(x)\big((I-\Delta_n)\chi_T\big)(x)\mu_n(x). \tag{32}$$

Alternatively, $E_{G_n}(S,T)$ was previously expressed (see [38]) as follows:

$$E_{G_n}(S,T) \quad = \quad n^2\int_{x\in S}\int_{y\in T} W(x,y)\ ds\ dt \tag{33}$$

The two formulations (32) and (33) look quite different but are of the same form when the graphs involved are regular. However, the format in (33) seems hard to extend to general graph sequences with smaller edge density.

Although the above definition in (31) seems complicated, it can be simplified when the degree sequences converge. Then, μ_m and μ_n are to be approximated by the measure μ of the graph limit. In such cases, we define

$$\begin{aligned}&\mathrm{disc}_\mu(G_m,G_n)\\ &= \inf_{\theta_m\in\mathcal{F}_m,\eta_n\in\mathcal{F}_n}\ \sup_{S,T\subseteq[0,1]}\left|\frac{\langle \chi_S,(I-\Delta_m)\chi_T\rangle_{\mu,\theta_m}}{\sqrt{\mu(S)\mu(T)}} - \frac{\langle \chi_S,(I-\Delta_n)\chi_T\rangle_{\mu,\eta_n}}{\sqrt{\mu(S)\mu(T)}}\right|\\ &= \sup_{S,T\subseteq[0,1]}\frac{1}{\sqrt{\mu(S)\mu(T)}}\left|\langle \chi_S,(\Delta_m-\Delta_n)\chi_T\rangle_\mu\right|.\end{aligned} \tag{34}$$

where S,T range over all integrable subsets of $[0,1]$ and we suppress the labelings θ,η which achieve the infininum.

We will show that the convergence using the spectral distance defined under the μ-norm is equivalent to the convergence using the discrepancy distance in Section 4.

4.2 The equivalence of convergence using spectral distance and the discrepancy distance

We will prove the following theorem concerning the equivalence of the convergences under the spectral distance (as in (19)) and the discrepancy distance (as in (31)). The result holds without any density restriction on the graph sequence. The proof extends similar techniques in Bilu and Linial [5] and [7, 11] for regular or random-like graphs to graph sequences of general degree distributions.

Theorem 2. *Suppose the degree distributions μ_n, of a graph sequence G_n, for $n = 1, 2, \dots$, converges to μ. The following statements are equivalent:*

(1) The graph sequence G_n converges under the spectral distance.

(2) The graph sequence G_n converges under the disc-*distance.*

Proof. Suppose that for a given $\epsilon > 0$, there exists an $N > 1/\epsilon$ such that for $n > N$, we have

$$\|\mu_n - \mu\|_1 < \epsilon.$$

The proof for (1) $\Rightarrow$ (2) is rather straightforward and can be shown as follows:

Suppose (1) holds and we have, for $m, n > N$, $\|\mu_m - \mu\|_1 < \epsilon$, $\|\mu_n - \mu\|_1 < \epsilon$ and $\|\Delta_m - \Delta_n\|_\mu < \epsilon$. (Here we omit the labeling maps θ_m, θ_n to simplify the notation.) Then,

$$\begin{aligned}
\text{disc}(G_m, G_n) &= \sup_{S,T\subseteq[0,1]} \left| \frac{\langle \chi_S, (I-\Delta_m)\chi_T\rangle_{\mu_m}}{\sqrt{\mu_m(S)\mu_m(T)}} - \frac{\langle \chi_S, (I-\Delta_n)\chi_T\rangle_{\mu_n}}{\sqrt{\mu_n(S)\mu_n(T)}} \right| \\
&\le \sup_{S,T\subseteq[0,1]} \frac{1}{\sqrt{\mu(S)\mu(T)}} |\langle \chi_S, (\Delta_m - \Delta_n)\chi_T\rangle_\mu| + 4\epsilon \\
&= \sup_{S,T\subseteq[0,1]} \frac{1}{\|\chi_S\|_\mu \|\chi_T\|_\mu} |\langle \chi_S, (\Delta_m - \Delta_n)\chi_T\rangle_\mu| + 4\epsilon \\
&\le \|\Delta_m - \Delta_n\|_\mu + 4\epsilon \\
&\le 5\epsilon.
\end{aligned}$$

To prove (2) $\Rightarrow$ (1), we assume that for $M = \Delta_n - \Delta_m$

$$|\langle \chi_S, M\chi_T\rangle_\mu| \le \epsilon\sqrt{\mu(S)\mu(T)} \tag{35}$$

for some $\epsilon > 0$ for any two integrable subsets $S, T \subseteq [0, 1]$. It is enough to show that for any two integrable functions $f, g : [0, 1] \to \mathbb{R}$, we have

$$|\langle f, Mg\rangle_\mu| \le 20\epsilon \log(1/\epsilon)\|f\|_\mu \|g\|_\mu \tag{36}$$

provided $\epsilon < .02$.

The proof of (36) follows a sequence of claims.

Claim 1: For an integrable function f defined on $[0,1]$ with $\|f\|_\mu = 1$, for any $\epsilon > 0$, there exists an $N(\epsilon)$ such that for any $n > N(\epsilon)$ there is a function h defined on $[0,1]$ satisfying :
(1) $\|h\|_\mu \leq 1$,
(2) $\|f-h\|_\mu \leq 1/4+\epsilon$,
(3) The value $h(y)$ in the interval $((j-1)/n, j/n]$ is a constant h_j and h_j is of the form $(\frac{4}{5})^j$ for integers j.
Proof of Claim 1: Since f is integrable, for a given ϵ, we can approximate $\|f\|_\mu^2$ by a function $\bar{f}$, with $\bar{f}(x) = f_j$ in $((j-1)/n, j/n]$, such that

$$\left|\int_0^1 (f-\bar{f})^2(x)\mu(x)\right| < \epsilon.$$

For $\bar{f} = (f_j)_{1\leq j\leq mn}$, we define $h = (h_j)_{1\leq j\leq mn}$ as follows. If $f_j = 0$, we set $h_j = 0$. Suppose $f_j \neq 0$, there is a unique integer k so that $(4/5)^k < |f_j| \leq (4/5)^{k-1}$. We set $h_j = \text{sign}(f)(\frac{4}{5})^k$ where $\text{sign}(f_j) = 1$ if f_j is positive and -1 otherwise. Then

$$0 < |f_j - h_j| \leq (\frac{4}{5})^{k-1} - (\frac{4}{5})^k = \frac{1}{4}(\frac{4}{5})^k < \frac{1}{4}|f_j|,$$

which implies $\|f-h\|_\mu^2 \leq \epsilon + \sum_j \int_0^1 |f_j - h_j|^2\mu(x) \leq \epsilon + \frac{1}{16}\sum_t |f_j|^2\mu(x) = \frac{1}{16}+\epsilon$. Claim 1 is proved.

Claim 2: Suppose there are functions f', g' satifying $\|M\|_\mu = |\langle f', Mg'\rangle_\mu|$ and $\|f'\|_\mu = \|g'\|_\mu = 1$. If f, g are functions such that $\|f\|_\mu, \|g\|_\mu \leq 1$ and $\|f' - f\|_\mu \leq 1/4+\epsilon$,$\|g'-g\|_\mu \leq 1/4+\epsilon$, then

$$\|M\|_\mu \leq (2+4\epsilon)|\langle f, Mg\rangle_\mu|. \tag{37}$$

Claim 2 can be proved by using Claim 1 as follows:

$$\begin{aligned}
\|M\|_\mu &= |\langle f', Mg'\rangle_\mu| \\
&\leq |\langle f, Mg\rangle_\mu| + |\langle f'-f, Mg\rangle_\mu| + |\langle f', M(g'-g)\rangle_\mu| \\
&\leq |\langle f, Mg\rangle_\mu| + \big(\frac{2}{4}+2\epsilon\big)\|M\|_\mu.
\end{aligned}$$

This implies $\|M\|_\mu \leq (2+4\epsilon)|\langle f, Mg\rangle_\mu|$, as desired.

From Claims 1 and 2, we can upper bound $\|M\|_\mu$ to within a multiplicative factor of $2+4\epsilon$ by bounding of $|\langle f, Mg\rangle_\mu|$ with f, g of the following form: Namely, $f = \sum_t (\frac{4}{5})^t f^{(t)}$, where the $f^{(t)}$ denotes the indicator function of $\{x : \bar{f}(x) = (\frac{4}{5})^t\}$. Similarly we write $g = \sum_t (\frac{4}{5})^t g^{(t)}$, where the $g^{(t)}$ denotes the indicator function of $\{y : \bar{g}(y) = (\frac{4}{5})^t\}$. Now we choose $\kappa = \log_{4/5} \epsilon$ and we consider

$$
\begin{aligned}
|\langle f, Mg\rangle_\mu| &\leq \sum_{s,t} (\frac{4}{5})^{s+t} \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&\leq \sum_{|s-t|\leq\kappa} (\frac{4}{5})^{s+t} \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&\quad + \sum_{s} (\frac{4}{5})^{2s+\kappa} \sum_{t} \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&\quad + \sum_{t} (\frac{4}{5})^{2t+\kappa} \sum_{s} \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&= X + Y + Z.
\end{aligned}
$$

We now bound the three terms separately. For a function f, we denote $\mu(f) = \mu(\mathrm{supp}(f))$ to be the measure of the support of f. Using the assumption (35) for $(0,1)$-vectors and the fact that $f^{(s)}$'s are orthogonal (as well as the $g^{(t)}$'s), we have

$$
\begin{aligned}
X &= \sum_{|s-t|\leq\kappa} (\frac{4}{5})^{s+t} \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&\leq \epsilon \sum_{|s-t|\leq\kappa} (\frac{4}{5})^{s+t} \sqrt{\mu(f^{(s)})\mu(g^{(t)})} \\
&\leq \frac{\epsilon}{2} \sum_{|s-t|\leq\kappa} \big((\frac{4}{5})^{2s} \mu(f^{(s)}) + (\frac{4}{5})^{2t} \mu(g^{(t)}) \big) \\
&\leq \frac{\epsilon(2\kappa+1)}{2} \big(\sum_{s} (\frac{4}{5})^{2s} \mu(f^{(s)}) + \sum_{t} (\frac{4}{5})^{2t} \mu(g^{(t)}) \big) \\
&\leq \epsilon(2\kappa+1),
\end{aligned}
$$

since each term can appear at most $2\kappa + 1$ times. For the second term we have, by using Lemmas 3, the following:

$$\begin{aligned}
Y &\le \sum_s (\frac{4}{5})^{2s+\kappa} \sum_t \left| \langle f^{(s)}, Mg^{(t)} \rangle_\mu \right| \\
&\le (\frac{4}{5})^\kappa \sum_s (\frac{4}{5})^{2s} \langle f^{(s)}, |(\Delta_m - \Delta_n) \sum_t g^{(t)}| \rangle_\mu \\
&\le (\frac{4}{5})^\kappa \sum_s (\frac{4}{5})^{2s} \langle f^{(s)}, (\Delta_m + \Delta_n)\mathbf{1} \rangle_\mu \\
&\le 2(\frac{4}{5})^\kappa \sum_s (\frac{4}{5})^{2s} \langle f^{(s)}, \mathbf{1} \rangle_\mu \\
&\le 2(\frac{4}{5})^\kappa \sum_s \mu(f^{(s)}) \\
&\le 2(\frac{4}{5})^\kappa
\end{aligned}$$

The third term can be bounded in a similar way. Together, we have

$$\begin{aligned}
\|M\|_\mu &\le (2+4\epsilon)\big(\epsilon(2\kappa+1) + 4(\frac{4}{5})^\kappa\big) \\
&\le (2+4\epsilon)\big(\epsilon(2\frac{\log(1/\epsilon)}{\log 5/4} + 1) + 4\epsilon\big) \\
&\le \frac{4+8\epsilon}{\log(5/4)} \epsilon \log(1/\epsilon) + 8\epsilon \\
&\le 20\epsilon \log(1/\epsilon)
\end{aligned}$$

since $\frac{4}{\log 5/4} \approx 17.93$ and $\epsilon < .02$. This completes the proof of the theorem. $\square$

5 Quasi-random graphlets with general degree distributions – graphlets of rank 1

We consider a graph sequence that consists of quasi-random graphs with degree distributions converging to some general degree distribution. We will give characterizations for a quasi-random graph sequence by stating a number of equivalent properties. Although the proof is mainly by summarizing previous known facts, the format of graph limits helps in simplifying the previous various statements for quasi-random graphs with general degree distributions including the cases for sparse graphs.

Theorem 3. *The following statements are equivalent for a graph sequence G_n, where $n = 1, 2, \ldots$.*

(i) *G_n's form a quasi-random sequence with degree distribution converging to μ.*

(*ii*) *The graph sequence* $G_n = (V_n, \Delta_n)$ *converges to the graphlets* $\mathcal{G} = (\Omega, \Delta)$ *where* Ω *is a measure space with measure* μ *and* $I - \Delta$ *is of rank* 1*, i.e.,* $I - \Delta$ *has one nontrivial eigenvalue* 1*. (Equivalently, for each* n*,* $I_n - \Delta_n$ *has all eigenvalue* $o(1)$ *with the exception of one eigenvalue* 1*.)*

(*iii*) *The graph sequence* $G_n = (V_n, \Delta_n)$ *converges to the graphlets* $\mathcal{G} = (\Omega, \Delta)$ *where* Ω *is a measure space with measure* μ *and the Laplace operator* Δ *on* Ω *satisfies*

$$\int_{x\in\Omega} f(x)\big((I-\Delta)g\big)(x)\mu(x) = \int_{x\in\Omega} f(x)\mu(x) \int_{x\in\Omega} g(x)\mu(x)$$

for any integrable $f, g : \Omega \to \mathbb{R}$*.*

(*iv*) *The degree distribution* μ_n *of* G_n *converges to* μ *and*

$$\|D_n^{-1/2}\big(A_n - \frac{D_n J D_n}{\text{vol}(G_n)}\big)D_n^{-1/2}\| = o(1)$$

where A_n *and* D_n *denote the adjacency matrix and diagonal degree matrix of* G_n*, respectively. Here* $\|\cdot\|$ *denotes the usual spectral norm (in* L_2*) and* J *denotes the all* 1*'s matrix.*

(*v*) *There exists a sequence* ϵ_n *which approaches* 0 *as* n *goes to infinity such that* G_n *satisfies the property* $P(\epsilon_n)$*, namely, that the degree distribution* μ_n *converges to* μ *and for all* $S, T \subseteq V_n$

$$P(\epsilon_n): \qquad \left|E(S,T) - \frac{\text{vol}(S)\text{vol}(T)}{\text{vol}(G_n)}\right| \le \epsilon_n \sqrt{\text{vol}(S)\text{vol}(T)} \tag{38}$$

where $E(S,T) = \sum_{s\in S, t\in T} A(s,t)$*.*

Remark 13. Before proceeding to prove Theorem 3, we note that a sequence of random graphs with degree distribution μ_n converging to μ is an example satisfying the above properties almost surely. Here we use random graph model $G_{\mathbf{d}}$ for a given degree sequence $\mathbf{d} = (d_v)_{v\in G}$ defined by choosing $\{u, v\}$ as an edge with probability $d_u d_v / \sum_s d_s$ for any two vertices u and v, (see [20]).

Remark 14. The above list of equivalent properties does not include the measurement of counting subgraphs. Indeed, the problem of enumerating subgraphs in a sparse graph can be inherently difficult because, for example, a random graph $G(n, p)$ with $p = o(n^{-1/2})$ contains very few four cycles. Consequently, the error bounds could be proportionally quite large.

Instead of counting C_4, we can consider an even cycle C_{2k} or the trace of $(2k)$th power, leading to the following condition:

(vi) For some constant k *(depending only on the degree sequence), a graph sequence* G_n *satisfies*

$$\big|\mathit{Trace}(I - \Delta_n)^k - 1\big| = o(1).$$

Remark 15. Suppose that in a graph G_n, all eigenvalues of $I - \Delta_n$ except for eigenvalue 1 are strictly smaller than 1. Then as k goes to infinity, the trace of the kth power of $I - \Delta$ approaches 1. How should (vi) be modified in a way that it can be an equivalent property to (i) through (v) ? We will leave this as an intriguing question.

Question 1. Is (vi) equivalent to (i) through (v) for some constant k depending only on Ω?

Remark 16. It is easily checked that (vi) implies (ii). For the case of dense graphs, the reverse direction holds [19]. For general graphs, to prove $(ii) \to (vi)$ involves the spectral distribution. For example, for a regular graph on n vertices and degree d, a necessary condition for (vi) to hold is that $nd^{k/2} \leq \epsilon_n$. In particular, if the spectrum of the graph satisfies the semi-circle law, then this necessarily condition is also sufficient. For a general graph, the necessary condition should be replaced by $n\bar{d}^{k/2} \leq \epsilon_n$ where $\bar{d}$ is the second order average degree, namely, $\bar{d} = \sum_v d_v^2 / \sum_v d_v$. Nevertheless, there are quasi-random graphs that satisfy (ii) but require k much larger than $2 \log n / \log \bar{d}$. For example, we can take the product of a quasi-random graph G_p and a complete graph K_q which is formed by replacing each vertex of G_p with a copy of K_q and replacing each edge in G_p by a complete bipartite graph $K_{q,q}$.

Question 2. A subgraph F is said to be *forcing* if when the number of occurrence of F in a graph G is close to (say, within a multiplicative factor of $1+\epsilon$) what is expected in a random graph with the same degree sequence, then all subgraphs with a bounded number k of vertices (where ϵ depends on k) occur in G close to the expected values in a random graph with the same degree sequence. A natural problem is to determine subgraphs which are forcing for quasi-random graphs with general degree sequences.

Proof of Theorem 3: We will show $(i) \Rightarrow (v) \Rightarrow (iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$.

We note that $(i) \Rightarrow (v)$ follows from the implications of quasi-randomness for graphs with general degree distributions [18]. Also, $(v) \Rightarrow (iv)$ follows from the fact that (iv) is one of the equivalent quasi-random properties.

To see $(iv) \Leftrightarrow (iii)$, we note that the Laplace operator Δ_n of G_n satisfies, for any $f, g : V(G_n) \to \mathbb{R}$,

$$\left| \int_x f(x)(I - \Delta_n) g(x) \mu_n(x) - \int_x f(x)\mu_n(x) \int_x g(x)\mu_n(x) \right|$$

$$= \quad |\langle f, (I - \Delta_n) g\rangle_{\mu_n} - \langle f, \mathbf{1}\rangle_{\mu_n} \langle g, \mathbf{1}\rangle_{\mu_n}|$$

$$= \quad \left| \sum_{u \in V_n} \frac{f(u) A_n g(u)}{\mathrm{vol}(G_n)} - \sum_{u \in V_n} f(u)\mu_n(u) \sum_{v \in V_n} g(v)\mu_n(v) \right|$$

$$= \quad \left| f' D_n^{-1/2} \left(A_n - \frac{D_n J D_n}{\mathrm{vol}(G_n)} \right) D_n^{-1/2} g' \right|$$

where $f' = D_n^{1/2} f/\mathrm{vol}(G_n)$ and $g' = D_n^{1/2} g/\mathrm{vol}(G_n)$. To prove $(iii) \Rightarrow (iv)$, we have from (iii),

$$\left| \int_x f(x)(I - \Delta) g(x) \mu_n(x) - \int_x f(x)\mu_n(x) \int_x g(x)\mu_n(x) \right|$$

$$\leq \quad \|D^{-1/2} \left(A_n - \frac{D_n J D_n}{\mathrm{vol}(G_n)} \right) D^{-1/2}\| \cdot \|f'\| \cdot \|g'\|$$

$$\leq \quad \epsilon_n \|f'\| \|g'\|$$

$$= \quad \epsilon_n \sqrt{\int f^2(x)\mu_n(x) \int g^2(x)\mu_n(x)}.$$

Since μ_n converges to μ and ϵ_n goes to 0 as n approaches infinity, $(iv) \Rightarrow (iii)$ is proved. The other direction can be proved in a similar way.

$(iii) \Rightarrow (ii)$ follows from the fact that $I - \Delta$ is of rank 1. All adjacency matrices A_n are close to a rank 1 matrix and therefore Ω is of rank 1.

To prove that $(ii) \Rightarrow (i)$, we use the fact that for any graph the Laplace operator is a sum of projections of eigenspaces. If Δ is of rank 1, there is only one main eigenspace of dimension 1 (associated with the Perron vector) for the normalized adjacency matrix. □

6 Bipartite quasi-random graphlets with general degree distributions

We consider the graph limit of a graph sequence consisting of bipartite quasi-random graphs with degree distributions converging to some general degree distribution. The characterizations for a bipartite quasi-random graph sequence are similar but different from those of quasi-random graphs. Because of the role that bipartite quasi-random graphlets plays in general graphlets, we will state a number of equivalent properties. The proof is quite similar to that for Theorem 3 and will be omitted.

Theorem 4. *The following statements are equivalent for a graph sequence G_n, where $n = 1, 2, \ldots$.*

(i) *G_n's form a bipartite quasi-random sequence with degree distribution converging to μ.*

(ii) *The graph sequence $G_n = (V_n, \Delta_n)$, converges to the graphlets $\mathcal{G} = (\Omega, \Delta)$ where Ω is a measure space with measure μ and and $I - \Delta$ has two nontrivial eigenvalues 1 and -1. Namely, for each n, $I - \Delta_n$ has all eigenvalues $o(1)$ with exceptions of two eigenvalues 1 and -1.*

(iii) *The graph sequence G_n converges to the graphlets $\mathcal{G} = (\Omega, \Delta)$ where Ω is a measure space with measure μ. For some $X \subset \Omega$, the Laplace operator Δ satisfies*

$$\begin{aligned}&\int_{x\in\Omega} f(x)(I-\Delta)g(x)\mu(x)\\&= \int_{x\in X} f(x)\mu(x)\int_{x\in\bar{X}} g(x)\mu(x) + \int_{x\in\bar{X}} f(x)\mu(x)\int_{x\in X} g(x)\mu(x)\end{aligned}$$

for any $f, g : \Omega \to \mathbb{R}$ where $\bar{X}$ denotes the complement of X.

(iv) *The degree distribution μ_n of G_n converges to μ and*

$$\|D_n^{-1/2}\big(A_n - \frac{D_n(J_{X,\bar{X}} + J_{\bar{X},X})D_n}{\mathrm{vol}(G_n)}\big)D_n^{-1/2}\| = o(1)$$

where $J_{X,\bar{X}}(x, y) = 1$ if ($x \in X$ and $y \in \bar{X}$) and 0 otherwise.

(v) *There exist $X \subset \Omega$ and a sequence ϵ_n which approaches 0 as n goes to infinity such that the bipartite graphs G_n satisfies the property that the degree distribution μ_n converges to μ and for all $S, T \subseteq V_n$*

$$\begin{aligned}&\left|E(S,T) - \frac{\big(\mathrm{vol}(S\cap X)\mathrm{vol}(T\cap\bar{X}) + \mathrm{vol}(S\cap\bar{X})\mathrm{vol}(T\cap X)\big)}{\mathrm{vol}(G_n)}\right|\\&\le \epsilon_n\sqrt{\mathrm{vol}(S)\mathrm{vol}(T)}\end{aligned}$$

where $E(S,T) = \sum_{s\in S, t\in T} A(s,t)$.

7 Graphlets with rank 2

It is quite natural to generalize rank 1 graphlets to graphlets of higher ranks. The case of rank 2 graphlets is particularly of interest, for example, in the sense for identifying two 'communities' in one massive graph. For two graphs with the same vertex set, the union of two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ has the edge set $E = E_1 \cup E_2$ and with edge weight $w(u,v) = w_1(u,v) + w_2(u,v)$ if w_i denotes the edge weights in G_i. We will prove the following theorem for graphlets of rank 2.

Theorem 5. *The following statements are equivalent for a graph sequence G_n, where $n = 1, 2, \ldots$. Here we assume that all G_n's are connected.*

(i) The graph sequence $G_n = (V_n, \Delta_n)$ converges to graphlets $\mathcal{G}(\Omega, \Delta)$ and $I - \Delta$ has two nontrivial eigenvalues 1 and $\rho \in (0,1)$. Namely, or each n, $I - \Delta_n$ has all eigenvalues $o(1)$ with the exception of two eigenvalue 1 and ρ_n where ρ_n converges to ρ.

(ii) The graph sequence G_n converges to the graphlets $\mathcal{G} = (\Omega, \Delta)$ which is the union of two quasi-random graphlets (of rank 1).

(iii) The graph sequence G_n converges to the graphlets $\mathcal{G} = (\Omega, \Delta)$ where Ω is a measure space with measure μ where $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ for some $\alpha \in [0,1]$ and the Laplace operator Δ on Ω satisfies

$$\begin{aligned}&\int_x f(x)(I-\Delta)g(x)\mu(x)\\ &= \alpha \int_\Omega f(x)\mu_1(x) \int_\Omega g(x)\mu_1(x) + (1-\alpha)\int_\Omega f(x)\mu_2(x)\int_\Omega g(x)\mu_2(x)\end{aligned}$$

for any $f, g : \Omega \to \mathbb{R}$.

(iv) The degree sequence $(d_v)_{v\in V}$ of G_n can be decomposed as $d_v = d'_v + d''_v$ with $d'_v \geq 0$ and $d''_v \geq 0$. The adjacency matrix A_n of G_n satisfies:

$$\|D_n^{-1/2}\big(A_n - \frac{D'_n J D'_n}{\mathrm{vol}(G'_n)} - \frac{D''_n J D''_n}{\mathrm{vol}(G''_n)}\big)D_n^{-1/2}\| = o(1)$$

where $\mathrm{vol}(G'_n) = \sum_v d'_v$ and $\mathrm{vol}(G''_n) = \sum_v d''_v$.

(v) There exists a sequence ϵ_n which approaches 0 as n goes to infinity such that the degree sequence $(d_v)_{v\in V}$ of G_n can be decomposed as $d_v = d'_v + d''_v$ with $d'_v \geq 0$ and $d''_v \geq 0$. Furthermore, for all $S, T \subseteq V_n$

$$\left| E_n(S,T) - \frac{\mathrm{vol}'(S)\mathrm{vol}'(T)}{\mathrm{vol}(G'_n)} - \frac{\mathrm{vol}''(S)\mathrm{vol}''(T)}{\mathrm{vol}(G''_n)} \right| \leq \epsilon_n \sqrt{\mathrm{vol}(S)\mathrm{vol}(T)}.$$

Before we proceed to prove Theorem 5, we first prove several key facts that will be used in the proof.

Lemma 4. *Suppose that integers d_v, d'_v and d''_v, for v in V satisfy $d_v = d'_v + d''_v$ and $d'_v, d''_v \geq 0$. Let D, D' and D'' denote the diagonal matrices with diagonal entries d_v, d'_v and d''_v, respectively. Then the matrix X defined by*

$$X = D^{-1/2} \left(\frac{D'JD'}{\mathrm{vol}(G')} + \frac{D''JD''}{\mathrm{vol}(G'')} \right) D^{-1/2}$$

has two nonzero eigenvalues 1 and η satisfying

$$\eta = 1 - \left(\sum_v \frac{d'_v d''_v}{d_v} \right) \left(\frac{\mathrm{vol}(G)}{\mathrm{vol}(G')\mathrm{vol}(G'')} \right).$$

The eigenvector ξ which is associated with eigenvalue η can be written as

$$\xi = D^{-1/2} \left(\frac{D'}{\mathrm{vol}(G')} - \frac{D''}{\mathrm{vol}(G'')} \right) \mathbf{1}.$$

Proof. The lemma will follow from the following two claims.

Claim 1: $\phi_0 = D^{1/2}\mathbf{1}/\sqrt{\mathrm{vol}(G)}$ is an eigenvector of X and $M = D^{-1/2}AD^{-1/2}$.
Proof of Claim 1: Following the definition of M, ϕ_0 is an eigenvector of M. We can directly verify that ϕ_0 is also an eigenvector of X as follows:

$$\begin{aligned} X\phi_0 &= D^{-1/2} \left(\frac{D'JD'}{\mathrm{vol}(G')} + \frac{D''JD''}{\mathrm{vol}(G'')} \right) \frac{\mathbf{1}}{\sqrt{\mathrm{vol}(G)}} \\ &= D^{-1/2} \left(D' + D'' \right) \frac{\mathbf{1}}{\sqrt{\mathrm{vol}(G)}} \\ &= D^{-1/2} \frac{D\mathbf{1}}{\sqrt{\mathrm{vol}(G)}} \\ &= \frac{D^{1/2}\mathbf{1}}{\sqrt{\mathrm{vol}(G)}}. \end{aligned}$$

Claim 2: η is an eigenvalue of X with the associated eigenvector ξ.

Proof of Claim 2: We consider

$$
\begin{aligned}
X\xi &= D^{-1/2}\left(\frac{D'JD'}{\mathrm{vol}(G')}+\frac{D''JD''}{\mathrm{vol}(G'')}\right)D^{-1}\left(\frac{D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\right)\\
&= D^{-1/2}\left(\frac{D'\mathbf{1}}{\mathrm{vol}(G')}\cdot\frac{\mathbf{1}^*D'D^{-1}D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{D'\mathbf{1}}{\mathrm{vol}(G')}\cdot\frac{\mathbf{1}^*D'D^{-1}D"\mathbf{1}}{\mathrm{vol}(G'')}\right)\\
&\quad + D^{-1/2}\left(\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\cdot\frac{\mathbf{1}^*D''D^{-1}D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\cdot\frac{\mathbf{1}^*D''D^{-1}D''\mathbf{1}}{\mathrm{vol}(G'')}\right)\\
&= D^{-1/2}\frac{D'\mathbf{1}}{\mathrm{vol}(G')}\left(\frac{\mathbf{1}^*D'D^{-1}(D-D'')\mathbf{1}}{\mathrm{vol}(G')}-\frac{\mathbf{1}^*D'D^{-1}D''\mathbf{1}}{\mathrm{vol}(G'')}\right)\\
&\quad + D^{-1/2}\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\left(\frac{\mathbf{1}^*D''D^{-1}D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{\mathbf{1}^*D''D^{-1}(D-D')\mathbf{1}}{\mathrm{vol}(G'')}\right)\\
&= D^{-1/2}\left(\frac{D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\right)\left(1-\mathbf{1}^*D'D^{-1}D''\mathbf{1}\left(\frac{1}{\mathrm{vol}(G')}+\frac{1}{\mathrm{vol}(G'')}\right)\right)\\
&= \eta\,\xi
\end{aligned}
$$

as claimed.

Since X has rank 2 (i.e., it is the sum of two rank one matrices), and we have shown that X has eigenvalues 1, η, then the rest of the eigenvalues are 0. □

We now apply Lemma 4 using the fact that the normalized adjacency matrix $M = D^{-1/2}AD^{-1/2}$ has eigenvalues 1 and $\rho = 1-\lambda_1$. Together with Theorem 3, we have the following:

Theorem 6. *Suppose G is the union of two graphs G' and $G''$with degree sequences (d'_v) and (d''_v) respectively. Assume both G' and G'' satisfy the quasi-random property $P(\epsilon/2)$ (where P is one of the equivalent quasi-random properties in Theorem 3). Suppose the normalized Laplacian of G has eigenvalues $\lambda_i = 1-\rho_i$, for $i = 0, 1, \ldots, n-1$ with associated orthonormal eigenvectors ϕ_i. Then we have:*

1. $\rho_0 = 1$,
2. ρ_1 *satisfies*
$$-\epsilon < 1-\rho_1-\left(\sum_v \frac{d'_v d''_v}{d_v}\right)\left(\frac{\mathrm{vol}(G)}{\mathrm{vol}(G')\mathrm{vol}(G'')}\right) < \epsilon$$
3. $|\rho_i| \le \epsilon$ *for* $i > 1$.
4. *The eigenvector ϕ_1 associated with λ_1 can be written as*
$$\phi_1 = D^{-1/2}\left(\frac{D'\mathbf{1}}{\mathrm{vol}(G')}-\frac{D''\mathbf{1}}{\mathrm{vol}(G'')}\right)+r$$

with $\|r\| \le \epsilon$*, where* D' *and* D'' *denote the diagonal degree matrices of* G' *and* G''*, respectively.*

Theorem 7. *Suppose a graphlets* $\mathcal{G} = (\Omega, \Delta)$ *is the union of two graphlets* $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ *and* $\mathcal{G}_i$ *are quasi-random graphlets. Then* $I - \Delta$ *has two nontrivial eigenvalues* 1 *and* η *where* $0 < \eta < 1$ *satisfies*

$$1 - \eta = \int_\Omega \frac{\mu_1(x)\mu_2(x)}{\mu(x)} = \langle \frac{\mu_1}{\mu}, \frac{\mu_2}{\mu} \rangle_\mu,$$

where μ_i *denotes the measure on* Ω_i*.*

Proof. The proof follows immediately from Lemma 4 by substituting $\mu_1(v) = d'(v)/\text{vol}(G')$ and $\mu_2(v) = d''(v)/\text{vol}(G'')$ in Lemma 4 and Theorem 6 before taking limit as n goes to infinity. □

In the other direction, we prove the following:

Theorem 8. *Suppose that the normalized adjacency matrix of a graph* G *has two nontrivial positive eigenvalues* 1 *and* ρ *and the other eigenvalues satisfy* $|\rho_i| \le \epsilon$ *for* $2 \le i \le n-1$*. Then for each vertex* v*, the degree* d_v *can be written as* $d_v = d'_v + d''_v$*, with* $d'_v, d''_v \ge 0$*, so that for any subset* S *of vertices, the number* $E(S)$ *of ordered pairs* (u, v)*, with* $u, v \in S$ *and* $\{u, v\} \in E$*, satisfies*

$$\left| E(S) - \frac{\text{vol}'(S)^2}{\text{vol}'(G)} - \frac{\text{vol}''(S)^2}{\text{vol}''(G)} \right| \le 2\epsilon \text{vol}(S)$$

where $\text{vol}'(S) = \sum_{v \in S} d'_v$ *and* $\text{vol}''(S) = \sum_{v \in S} d''_v$*.*

Proof. Let ϕ_i, $0 \le i \le n-1$, denote the eigenvectors of the normalized adjacency matrix of G. Let ϕ_0 and ϕ_1 denote the eigenfunctions associated with $\rho_0 = 1$ and ρ_1.

Since G is connected, the eigenvector ϕ_0 associated with eigenvalue $\rho_0 = 1$ of M_G can be written as $\phi_0 = D^{1/2}\mathbf{1}/\sqrt{\text{vol}(G)}$ as seen in [13]. The second largest eigenvalue ρ_1 is strictly between 0 and 1 because of the connectivity of G. Before we proceed to analyze the eigenvector ϕ_1 associated with ρ_1, we consider the following two vectors which depend on a value α to be specified later.

$$\begin{aligned} f_1 &= \alpha D\mathbf{1} - D^{1/2}\phi_1\sqrt{\rho_1\alpha(1-\alpha)\text{vol}(G)} \\ f_2 &= (1-\alpha)D\mathbf{1} + D^{1/2}\phi_1\sqrt{\rho_1\alpha(1-\alpha)\text{vol}(G)} \end{aligned} \tag{39}$$

It is easy to verify that f_1 and f_2 satisfy the following:

$$f_1 + f_2 = D\mathbf{1} \tag{40}$$

$$\mathbf{1} \perp \left(\frac{f_1}{\alpha} - \frac{f_2}{1-\alpha}\right) \tag{41}$$

$$\sum_v f_1(v) = \alpha \mathrm{vol}(G),$$

$$\sum_v f_2(v) = (1-\alpha)\mathrm{vol}(G).$$

In particular, by considering $\langle f_1, D^{-1} f_2 \rangle$, we see that α satisfies

$$1 - \rho_1 = \frac{1}{\alpha(1-\alpha)\mathrm{vol}(G)} \sum_v \frac{f_1(v) f_2(v)}{d_v}. \tag{42}$$

and we have

$$\phi_1 = \sqrt{\frac{\alpha(1-\alpha)}{\rho_1 \mathrm{vol}(G)}} D^{-1/2} \left(\frac{f_1}{\alpha} - \frac{f_2}{1-\alpha}\right).$$

Claim A:

$$\phi_0 \phi_0^* + \rho_1 \phi_1 \phi_1^* = \frac{D^{-1/2} f_1 f_1^* D^{-1/2}}{\alpha \mathrm{vol}(G)} + \frac{D^{-1/2} f_2 f_2^* D^{-1/2}}{(1-\alpha)\mathrm{vol}(G)}.$$

Proof of Claim A:
From (39), we have

$$\frac{D^{-1/2} f_1 f_1^* D^{-1/2}}{\alpha \mathrm{vol}(G)} = \alpha \frac{D^{1/2} J D^{1/2}}{\mathrm{vol}(G)} + (1-\alpha)\rho_1 \phi_1 \phi_1^*.$$

Similarly, we have

$$\frac{D^{-1/2} f_2 f_2^* D^{-1/2}}{(1-\alpha)\mathrm{vol}(G)} = (1-\alpha) \frac{D^{1/2} J D^{1/2}}{\mathrm{vol}(G)} + \alpha \rho_1 \phi_1 \phi_1^*.$$

Combining the above two equalities, Claim A is proved.

Now, we define two subsets X and Y satisfying

$$X = \{x : f_1(x) < 0\} = \left\{x : d_x^{1/2} \leq \phi_1(x) \sqrt{\frac{(1-\alpha)\mathrm{vol}(G)}{\alpha}}\right\}$$

$$Y = \{y : f_2(y) < 0\} = \left\{y : d_y^{1/2} < -\phi_1(y) \sqrt{\frac{\alpha \mathrm{vol}(G)}{1-\alpha}}\right\}.$$

Clearly X and Y are disjoint.

Note that when α decreases, the volume of X decreases and the volume of Y increases. If $\alpha = 1$, X consists of all v with $\phi_1(v) \geq 0$ and Y is empty. For $\alpha = 0$, Y consists of all u with $\phi_1(u) < 0$ and X is empty. We choose α so that

$$\sum_{x \in X} |f_1(x)| = \sum_{y \in Y} |f_2(y)|. \tag{43}$$

Here we use the convention that a subset X' of X means that there are values γ_v in $\{0,1\}$, associated each vertex in X with the exception of one vertex with a fractional γ_v and the size of X' is the sum of all γ_vs.

Now, for each vertex v, we define d'_v and d''_v as follows:

$$d'_v = \begin{cases} f_1(v) & \text{if } v \notin X \cup Y, \\ 0 & \text{if } v \in X, \\ d_v & \text{if } v \in Y. \end{cases} \tag{44}$$

Also, we define $d''_v = d_v - d'_v$.

Claim B:

$$\sum_v d'_v = \alpha \mathrm{vol}(G)$$
$$\sum_v d''_v = (1-\alpha)\mathrm{vol}(G)$$

Proof of Claim B: We note that

$$\begin{aligned}
\sum_v d'_v - \alpha \mathrm{vol}(G) &= \sum_v d'(v) - \sum_v f_1(v) \\
&= \sum_{v \in X \cup Y} (d'_v - f_1(v)) \\
&= \sum_{x \in X} |f_1(x)| + \sum_{y \in Y} (d_y - f_1(y)) \\
&= \sum_{x \in X} |f_1(x)| + \sum_{y \in Y} f_2(y) \\
&= \sum_{x \in X} |f_1(x)| - \sum_{y \in Y} |f_2(y)| \\
&= 0.
\end{aligned}$$

The second equality can be proved in a similar way that completes the proof of Claim B.

For a subset S of vertices, let χ_S denote the characteristic function of S

defined by $\chi_S(x) = 1$ if x in S and 0 otherwise. We consider

$$\begin{aligned}0 &\le \chi_X^* D^{1/2} M D^{1/2} \chi_Y \\ &\le \chi_X^* D^{1/2} (\phi_0 \phi_0^* + \rho_1 \phi_1 \phi_1^*) D^{1/2} \chi_Y + \epsilon \|D^{1/2}\chi_X\| \, \|D^{1/2}\chi_Y\| \\ &= \frac{\chi_X^* f_1 f_1^* \chi_Y}{\alpha \mathrm{vol}(G)} + \frac{\chi_X^* f_2 f_2^* \chi_Y}{(1-\alpha)\mathrm{vol}(G)} + \epsilon \sqrt{\mathrm{vol}(X)\mathrm{vol}(Y)}. \end{aligned} \tag{45}$$

From the definition, we have $\chi_X^* f_1 < 0$, $\chi_Y^* f_1 > 0$, $\chi_X^* f_2 > 0$ and $\chi_Y^* f_2 < 0$. This implies

$$\begin{aligned}\epsilon\sqrt{\mathrm{vol}(X)\mathrm{vol}(Y)} &\ge -\frac{\chi_X^* f_1 f_1^* \chi_Y}{\alpha \mathrm{vol}(G)} - \frac{\chi_X^* f_2 f_2^* \chi_Y}{(1-\alpha)\mathrm{vol}(G)} \\ &= \left| \frac{\chi_X^* f_1 f_1^* \chi_Y}{\alpha \mathrm{vol}(G)} \right| + \left| \frac{\chi_X^* f_2 f_2^* \chi_Y}{(1-\alpha)\mathrm{vol}(G)} \right| \\ &= \frac{|f_1^* \chi_X|(\mathrm{vol}(Y) - f_2^* \chi_Y)}{\alpha \mathrm{vol}(G)} + \frac{|f_2^* \chi_Y|(\mathrm{vol}(X) - f_1^* \chi_X)}{(1-\alpha)\mathrm{vol}(G)} \\ &= \frac{|f_1^* \chi_X|(\mathrm{vol}(Y) + |f_2^* \chi_Y|)}{\alpha \mathrm{vol}(G)} + \frac{|f_2^* \chi_Y|(\mathrm{vol}(X) + |f_1^* \chi_X|)}{(1-\alpha)\mathrm{vol}(G)} \\ &\ge \frac{|f_1^* \chi_X|}{\mathrm{vol}(G)} \left(\frac{\mathrm{vol}(Y)}{\alpha} + \frac{\mathrm{vol}(X)}{1-\alpha}\right) \end{aligned} \tag{46}$$

by using (39) and (40). Now, we have

$$\begin{aligned}\frac{\mathrm{vol}(Y)}{\alpha} + \frac{\mathrm{vol}(X)}{1-\alpha} &= \alpha \left(\frac{\sqrt{\mathrm{vol}(Y)}}{\alpha} \right)^2 + (1-\alpha) \left(\frac{\sqrt{\mathrm{vol}(X)}}{1-\alpha} \right)^2 \\ &\ge \left(\sqrt{\mathrm{vol}(X)} + \sqrt{\mathrm{vol}(Y)} \right)^2 \\ &\ge 4\sqrt{\mathrm{vol}(X)\mathrm{vol}(Y)} \end{aligned} \tag{47}$$

by using the Cauchy-Schwarz inequality. Combining (46) and (47), we have

$$|f_1^* \chi_X| = |f_2^* \chi_Y| \le \frac{\epsilon}{4} \mathrm{vol}(G). \tag{48}$$

Now we consider

$$R = A - \frac{D' J D'}{\sum_v d'_v} - \frac{D'' J D''}{\sum_v d''_v}.$$

Then, for $f = \chi_S$, the characteristic function of the subset S, we have

$$\begin{aligned}
\langle f, Rf \rangle &= f^* D^{1/2} M D^{1/2} f - \frac{f^* D' J D' f}{\sum_v d'_v} - \frac{f^* D'' J D'' f}{\sum_v d''_v} \\
&\le f^* D^{1/2} (\phi_0 \phi_0^* + \rho_1 \phi_1 \phi_1^*) D^{1/2} f \\
&\quad - \frac{f^* D' J D' f}{\sum_v d'_v} - \frac{f^* D'' J D'' f}{\sum_v d''_v} + 2\epsilon \|D^{1/2} f\|^2 \\
&\le \frac{f^* f_1 f_1^* f}{\alpha \mathrm{vol}(G)} + \frac{f^* f_2 f_2^* f}{(1-\alpha)\mathrm{vol}(G)} - \frac{f^* D' J D' f}{\sum_v d'_v} - \frac{f^* D'' J D'' f}{\sum_v d''_v} + 2\epsilon \mathrm{vol}(S) \\
&\le \frac{(f^* f_1)^2 - (f^* \mathbf{d}')^2}{\alpha \mathrm{vol}(G)} + \frac{(f^* f_2)^2 - (f^* \mathbf{d}'')^2}{(1-\alpha)\mathrm{vol}(G)} + 2\epsilon \mathrm{vol}(S).
\end{aligned}$$

where $\mathbf{d}'$ and $\mathbf{d}''$ are the degree vectors with entries d'_v and d''_v, respectively.

Since $f = \chi_S$, we have

$$\begin{aligned}
\frac{(f^* f_1)^2 - (f^* \mathbf{d}')^2}{\alpha \mathrm{vol}(G)} &\le \frac{2\sum_{v \in S \cap X} |f_1(v)| \mathrm{vol}'(S) + \sum_{v \in S \cap X} |f_1(v)|^2}{\alpha \mathrm{vol}(G)} \\
&\le 3\epsilon \mathrm{vol}(S)
\end{aligned}$$

Similar inequalities hold for f_2 and $\mathbf{d}''$. Thus, we have

$$\langle f, Rf \rangle \le 8\epsilon \mathrm{vol}(S)$$

The proof of Theorem 8 is complete. □

Theorem 9. *Suppose $\mathcal{G} = (\Omega, \Delta)$ is a graphlets and $I - \Delta$ has two nontrivial eigenvalues* 1 *and ρ with $0 < \rho < 1$. Then there is a value $\alpha \in [0, 1]$ such that*
(i) $\Omega = \Omega_1 \cup \Omega_2$ where $\mu(\Omega_1) = \alpha$ and $\mu(\Omega_2) = 1 - \alpha$,
(ii) Ω_i has a measure μ_i satisfying

$$\begin{aligned}
\mu_1(x) &= \mu(x) + \sqrt{\frac{\alpha\rho}{1-\alpha}} \mu(x) \varphi_1(x), \\
\mu_2(x) &= \mu(x) - \sqrt{\frac{(1-\alpha)\rho}{\alpha}} \mu(x) \varphi_1(x),
\end{aligned}$$

where φ_1 is the eigenvector, with $\|\varphi_1\|_\mu = 1$, associated with ρ.

The proof of Theorem 9 follows from the proof in Theorem 8 and Lemma 4. Thus, we have $(i) \Leftrightarrow (ii)$.

Proof of Theorem 5: We note that in the statement of Theorem 5, the implications $(ii) \Leftrightarrow (iv) \Leftrightarrow (v)$ follow from the definitions and Lemma 4. It suffices to prove $(i) \Leftrightarrow (ii)$ and $(iii) \Leftrightarrow (iv)$.

The implication $(ii) \Rightarrow (i)$ is proved in Theorem 6, and Theorems 8 and 9 implies $(i) \Rightarrow (ii)$.

To see that $(iii) \Leftrightarrow (iv)$, we note that if in a graph G_n in the graph sequence, the degree sequence d_x can be written as $d_v = d'_v + d''_v$ for all $v \in V(G_n)$ where $d'_v, d''_v \geq 0$, then by defining $\mu_1^{(n)}(v) = d'_v / \sum_v d'_v$, $\mu_2^{(n)}(v) = d''_v / \sum_v d''_v$ and $\alpha = \sum_v d'_v / \sum_v d_v$, we have $\mu_n = \mu_1^{(n)} + \mu_2^{(n)}$. Furthermore, we can use the fact that

$$\int_x f(x)(I-\Delta_n)g(x)\mu_n(x) = \frac{1}{\mathrm{vol}(G_n)}\langle f, (I-\Delta_n)g\rangle_{\mu_n}$$

and

$$\begin{aligned}\langle f, \mathbf{1}\rangle_{\mu_1^{(n)}}\langle g, \mathbf{1}\rangle_{\mu_1^{(n)}} &= \sum_{u\in V_n} f(u)\mu_1^{(n)}(u)\sum_{v\in V_n} g(v)\mu_2^{(n)}(v) \\ &= \frac{fD'_n J D''_n g}{\mathrm{vol}(G_n)^2}.\end{aligned}$$

The equivalence of (iii) and (iv) follows from substitutions using the above two equations and applying Theorem 3. Theorem 5 is proved. $\square$

For graphlets of rank 2, there can be a negative eigenvalue $-\rho$ of $I-\Delta$ in addition to the eigenvalue 1. For example, bipartite quasi-random graphlets have eigenvalues 1 and -1 for $I-\Delta$. In general, can such graphlets be characterized as the union of a quasi-random graphlet and a bipartite quasi-random graphlets? To this question, the answer is negative. It is not hard to construct examples of a graphlets having three nontrivial eigenvalues which is the union of a quasi-random graphlets and a bipartite quasi-random graphlets. With additional restrictions on degree distributions and edge density, the three eigenvalues can collapse into two eigenvalues. It is possible to apply similar methods as in the proof of Theorem 8 to derive the necessary and sufficient conditions for such cases but we will not delve into the details here.

8 Graphlets of rank k

In this section, we examine graphlets of rank k for some given positive integer k. It would be desirable to derive some general characterizations for graphlets of rank k, for example, similar to Theorem 5. However, for $k \geq 3$, the situation is more complicated. Some of the methods for the case of $k = 2$ can be extended but some techniques in the proof of Theorem 5 do not. Here we state a few useful facts about graphlets of rank k and leave some discussion in the last section.

Lemma 5. *Suppose D is the diagonal degree matrix of a graph G. Suppose that for all v in V, $d_v = \sum_{i=1}^k d_i(v)$, for $d_i(v) \geq 0$, $1 \leq i \leq k$. Let D_i denote the diagonal matrices with diagonal entries $D_i(v,v) = d_i(v)$. Then the matrix X*

defined by

$$X = D^{-1/2}\left(\sum_{i=1}^{k} \frac{D_i J D_i}{\mathrm{vol}_i(G)}\right) D^{-1/2}$$

has k nonzero eigenvalues η_i where η_i are eigenvalues of a $k \times k$ matrix M defined by

$$M(i,j) = \sum_v \frac{d_i(v)d_j(v)}{d_v}.$$

Furthermore, the eigenvector ξ_i for X which is associated with eigenvalue η_i can be written as

$$\xi_i(v) = \sum_{j=1}^{k} \psi_i(j) \frac{d_j(v)d_v^{-1/2}}{\mathrm{vol}_j(G)}$$

where ψ_i are eigenvectors of M associated with eigenvalues η_i.

Proof. The proof of Lemma 5 is by straightforward verification. Under the assumption that $\varphi_j M = \eta_j \varphi_i$ for $1 \le i \le k$, it suffices to check that $\xi_i X = \eta_i \xi$ for ξ_i. The proof is done by direct substitution and will be omitted. □

Theorem 10. *If a graphlets $\mathcal{G}(\Omega, \Delta)$ is the union of k quasi-random graphlets, then the Laplace operator Δ satisfies the property that $I - \Delta$ has k nontrivial positive eigenvalues.*

The proof of Theorem 10 follows immediately from Lemma 5.

Several questions follow the above theorem. If $I - \Delta$ has k eigenvalues that are not necessarily positive, is it possible to find a decomposition into a number of quasi-random graphlets or bipartite quasi-random graphlets? Under what additional conditions can such decompositions exist? If they exist, are they unique? Numerous additional questions can be asked here.

9 Concluding remarks

In this paper, we have merely scatched the surface of the study of graphlets. Numerous questions remain, some of which we mention here.

(1) In this paper, we mainly study quasi-random graphlets and graphlets of finite rank (which are basically 'sums' of quasi-random graphlets). It will be quite essential to understand other families of graph sequences, such as the graph sequences of paths, cycles, trees, grids, planar graphs, etc. In this paper, we define the spectral distance between two graphs as the spectral norm of the 'difference' of the associated Laplacians. In a subsequent paper, we consider a generalized version of spectral distance for considering large families of graphlets.

(2) We here use $[0,1]$ as the labels for the graphlets and the measure μ of the graphlets depends on the Lebesgue measure on $[0,1]$. To fully understand the geometry of graphlets derived from general graph sequences, it seems essential to consider general measurable spaces as labeling spaces. For example, for graph sequences $C_n \times C_n$, it works better to use $[0,1] \times [0,1]$ as the labeling space, instead.

(3) In this paper we relate the spectral distance to the previously studied cut-distance by showing the equivalence of the two distance measures for graph sequences of any degree distribution. It will be of interest to find and to relate to other distances. For example, will some nontrivial subgraph count measures be implied by the spectral distance (see the questions and remarks mentioned in Section 5)?

(4) In the study of complex graphs motivated by numerous real-world networks, random graphs are often utilized for analyzing various models of networks. Instead of using the classical Erdős-Rényi model, for which graphs have the same expected degree for every vertex, the graphs under consideration usually have prescribed degree distributions, such as a power law degree distribution. For example, for a given expected degree sequence $\mathbf{w} = (d_v)$, for $v \in V$, a random graph $G(\mathbf{w})$ has edges between u and v with probability $pd_u d_v$, for some scaling constant (see [20]). Such random graphs are basically quasi-random of positive rank one. Nevertheless, realistic networks often are clustered or have uneven distributions. A natural problem of interest is to identify the clusters or 'local communities'. The study of graphlets of rank two or higher can be regarded as extensions of the previous models. Indeed, the geometry of the graphlets can be used to illustrate the limiting behavior of large complex networks. In the other direction, network properties that are ubiquitous in many examples of real-world graphs can be a rich source for new directions in graphlets.

(5) Although we consider undirected graphs here, some of these questions can be extended to directed graphs. In this paper, we focus on the spectral distance of graphs but for directed graphs the spectral gaps can be exponentially small and any diffusion process on directed graphs can have very different behavior. The treatment for directed graphs will need to take these considerations into account. Many questions remains.

References

[1] D.J. Aldous, Representations for partially exchangeable arrays of random variables. J. Multivariate Anal. **11** (4) (1981) 581–598.

[2] D.J. Aldous, More uses of exchangeability: representations of complex random structures. In N.H. Bingham and C.M. Goldie, editors, Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman, Cambridge University Press, Cambridge, 2009.

[3] I. Benjamini, O. Schramm, Recurrence of distributional limits of finite planar graphs, Electron. J. Probab. **6** (23), (2001) 13 pp. (electronic).

[4] P. Bérard, G. Besson, S. Gallot, Embedding Riemannian manifolds by their heat kernel, Geom. Funct. Anal. **4** (4), (1994), 373–398.

[5] Y. Bilu, N. Linial, Constructing Expander Graphs by 2-Lifts and Discrepancy vs. Spectral Gap. Annual IEEE Symposium on Foundations of Computer Science, 45th Annual IEEE Symposium on Foundations of Computer Science, (FOCS'04), (2004) 404–412.

[6] B. Bollobás, S. Janson, O. Riordan, Monotone graph limits and quasimonotone graphs, Internet Math. **8** (2012), 187–231.

[7] B. Bollobás, V. Nikiforov, Hermitian matrices and graphs: singular values and discrepancy, Discrete Math. **285** (2004), 17–32.

[8] B. Bollobás, O. Riordan, Sparse graphs: Metrics and random models, Random Structures Algorithms, **39** (2011), 1–38.

[9] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi: Convergent Sequences of Dense Graphs I: Subgraph Frequencies, Metric Properties and Testing, Adv. Math. 219 (2008), 1801–1851.

[10] C. Borgs, J. Chayes, and L. Lovász. Moments of two-variable functions and the uniqueness of graph limits, Geom. Funct. Anal. **19** (6), (2010), 1597–1619.

[11] S. Butler, Using discrepancy to control singular values for nonnegative matrices, Linear Algebra Appl. **419** (2006), 486–493.

[12] H. Chernoff, A note on an inequality involving the normal distribution, Ann. Probab. **9** (1981), 533–535.

[13] F. Chung, Spectral Graph Theory, Amer. Math. Soc., Providence RI, 1997.

[14] F. Chung, The heat kernel as the pagerank of a graph, Proc. Nat. Acad. Sci. USA, **105** (50), (2007), 19735–19740.

[15] F. Chung, A local partitioning algorithm using heat kernel pagerank, WAW 2009, Lecture Notes in Computer Science 5427 (2009), 62–75.

[16] F. R. K. Chung, R. L. Graham, Quasi-random set systems, J. Amer. Math. Soc. **4** (1991), 151-196.

[17] F. Chung, R. L. Graham, Sparse quasi-random graphs, Combinatorica, **22** (2002), 217–244.

[18] F. Chung, R. L. Graham, Quasi-random graphs with given degree sequences, Random Structures Algorithms, **12** (2008), 1–19.

[19] F. R. K. Chung, R. L. Graham, R. M. Wilson, Quasi-random graphs, Combinatorica **9** (1989), 345-362.

[20] F. Chung, L. Lu, Complex Graphs and Networks, CBMS 107, Amer. Math. Soc., Providence, RI, 2006.

[21] F. Chung, L. Lu, V. Vu, The spectra of random graphs with given expected degrees, Proc. Nat. Acad. Sci. USA, **100** (11), (2003), 6293–6298.

[22] F. Chung, M. Radcliff, On the spectra of general random graphs, Electronic J. Combin. **18** (1), (2011), P215.

[23] F. Chung, S.-T. Yau, Coverings, heat kernels and spanning trees, Electron. J. Combin. **6** (1999) #R12.

[24] P. Diaconis , S. Janson, Graph limits and exchangeable random graphs, Rend. Mat. Appl. **28** (2008) 3361.

[25] P. Diaconis, S. Holmes and S. Janson, Threshold graph limits and random threshold graphs, Internet Math. **5** (3), (2008), 267–298.

[26] G. Elek, On limits of finite graphs, Combinatorica **27** (4), (2007), 503–507.

[27] G. Elek, L^2-spectral invariants and convergent sequences of finite graphs, J. Funct. Anal. **254** (2008) 2667–2689.

[28] B. de Finetti, La probabilitá e la statistica nei rapporti con induzione; secondo i diversi punti di vista, In Corso C. I. M. E. su Induzione e Statistica, Rome, 1959, (translated in Chapter 9 of de Finetti 1974).

[29] B. de Finetti, Probability, Induction, and Statistics, Wiley, New York, 1974.

[30] A. Frieze, R. Kannan, Quick approximation to matrices and applications, Combinatorica **19** (1999), 175–220.

[31] M. Gromov, Filling Riemannian manifolds, J. Differential Geom. **18** (1983), 1–147.

[32] H. Hatami, Graph norms and Sidoranko's conjecture, Israel J. Math. **175** (1) (2010), 125–150.

[33] H. Hatami, S. Norine, Undecidability of linear inequalities in graph homomorphism densities, J. Amer. Math. Soc. **24** (2), (2011) 547–565.

[34] D. Hoover, Relations on Probability Spaces and arrays of random variables, preprint, Institute for Advanced Study, Princeton, NJ, 1979.

[35] P. Li, S. T. Yau, Estimates of eigenvalues of a compact Riemannian manifold, Amer. Math. Soc. Proc., Symp. Pure Math. **36** (1980), 205–240.

[36] L. Lovász, Large networks and graph limits, Amer. Math. Soc., Colloquium Publications, Vol. 60, 2012.

[37] L. Lovász, V.T. Sós, Generalized quasirandom graphs, J. Combin. Theory, Ser. B, **98** (2008), 146–163.

[38] L. Lovász, B. Szegedy, Limits of dense graph sequences, J. Combin. Theory, Ser. B, **96** (2006), 933957.

[39] L. Lovász, B. Szegedy, Regularity partitions and the topology of graphons, in “An Irregular Mind (Szemerédi is 70)”, Bolyai Soc., Math. Studies, **21** (2010), 415–446

[40] L. Lovász, B. Szegedy, Finitely forcible graphons, J. Combin. Theory, Ser. B, **101** (2011), 269–301.

[41] L. Lovász, B. Szegedy, Testing properties of graphs and functions, Israel J. Math. **178** (2010), 113–156.

[42] L. Lovász, B. Szegedy, The graph theoretic moment problem, preprint.

[43] B. Szegedy, Limits of kernel operators and the spectral regularity lemma, European J. Combin. **32** (2011), 1156–1167.

[44] E. Szemerédi, Regular partitions of graphs, Colloques Internationaux, C. N. R. S. No. 260, Problémes Combinatoires et Théorie des Graphes, Orsay (1976), 399–401.

[45] S. Young, personal communications.