Limits for embedding distributions

Yichao Chen
chengraph@163.com

With Jinlian Zhang

Suzhou University of Science and Technology, Suzhou, China
By the **genus distribution** of a graph $G$, we mean the sequence

$$
\gamma_0(G), \gamma_1(G), \gamma_2(G), \ldots,
$$

where $\gamma_i(G)$ is the number of **distinct embeddings** of $G$ with genus $i$ for $i \geq 0$.

The **genus polynomial** of $G$ is $\Gamma_G(x) = \sum_{k=0}^{\infty} \gamma_k(G)x^k$. 
For any graph $G$, let $X_G$ be a random variable with distribution

$$p_i = P(X_G = i) = \frac{\gamma_i(G)}{\Gamma_G(1)}, \quad i = 0, 1, \cdots. \quad (1)$$

The probability genus polynomial of $G$ is defined as

$$P_{X_G}(z) = \sum_{i \geq 0} p_i z^i.$$
The crosscap-number distribution:
\[ \tilde{\gamma}_1(G), \tilde{\gamma}_2(G), \ldots \]

The crosscap-number polynomial:
\[ \tilde{\Gamma}_G(x) = \sum_{j=1}^{\infty} \tilde{\gamma}_j(G)x^j. \]

The Euler-genus distribution:
\[ \varepsilon_0(G), \varepsilon_1(G), \varepsilon_2(G), \ldots \]

The Euler-genus polynomial:
\[ \varepsilon_G(x) = \sum_{i=0}^{\infty} \varepsilon_i(G)x^i. \]

Similarly, we have the probability Euler-genus (crosscap-number) polynomial of \( G \).
For any graph $G$, it holds that

$$
\epsilon_G(x) = \Gamma_G(x^2) + \tilde{\Gamma}_G(x).
$$

When we say embedding distribution of a graph $G$, we mean its genus distribution, crosscap-number distribution or Euler-genus distribution.
Global feature for embedding distribution


- Partial results obtained by Gross and his coauthors, Stahl (1997) et. al.
Mean of embedding distribution or Average genus, average crosscap-number and average Euler-genus.

Limit for probability embedding distribution $X_n$.

- **Problem**: when $n$ is big enough, whether the distribution $X_n$ will converge to some well-known distribution in probability.

- If the answer is yes, then it demonstrates the outline of embedding distribution for graph $G_n$ when $n$ is big enough.
Suppose \( \{G_n\}_{n=1}^{\infty} \) is a sequence of graphs. For \( n \geq 1 \), let \( e_n \) and \( \sigma_n \) be the mean and variance of embedding distribution of \( G_n \), respectively. We say the embedding distribution of \( G_n \) is asymptotically normal distribution when \( n \) tends to infinity if for any \( x \in \mathbb{R} \), we have

\[
\lim_{n \to \infty} \mathbb{P}\left( \frac{X_{G_n} - e_n}{\sqrt{\sigma_n}} \leq x \right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \, du.
\]
Two interesting properties of normal distributions:

1. **Symmetry**, normal distributions are symmetric around their mean.
2. Approximately 95% of the area of a normal distribution is within two standard deviations of the mean.
If the embedding distribution of $G_n$ is asymptotically normal distribution, then the number of embeddings of $G_n$ are mainly concentrated on the interval $(e_n - 2\sqrt{v_n}, e_n + 2\sqrt{v_n})$ when $n$ is big enough.
The genus distribution of ladders was obtained by Furst, Gross and Stateman (1987).

Let $c > 0$ be a constant. We divide the interval $(-c, c]$ into $m$ intervals

$$I_i = (u_{i-1}, u_i] := (−c + 2(i−1)c/m, −c + 2ic/m],$$
Numeric Simulations for ladder graphs

- $m = 7, n = 13199$, and $m = 7, n = 14199$
A sequence \( \{G_n\}_{n=1}^{\infty} \) of graphs is called **strictly monotone sequence** if no pair of graphs in the sequence are homeomorphic and each \( G_i \) is homeomorphic to a subgraph of \( G_{i+1} \) for all \( i > 1 \).

We say a strictly monotone graph sequence \( \{G_n\}_{n=1}^{\infty} \) **bounded** if there exists a positive constant \( C \) such that \( \gamma_{max}(G_i) \leq C \) for \( i \geq 1 \).
A central limit theorem for \( \{G_n\}_{n=1}^{\infty} \)

**Theorem (B)**

Let \( \{G_n\}_{n=1}^{\infty} \) be a strictly monotone sequence of connected graphs. If \( \{G_n\}_{n=1}^{\infty} \) is bounded, then the Euler-genus distribution of \( G_n \) is asymptotically normal distribution.
Sketch of proof. Since the strictly monotone sequence of connected graphs $G_1, G_2, G_3, \cdots$, is bounded, then the values of the maximum genus of the graphs approach a finite limit point, and there exists an index $N$ such that all but a finite number of graphs in the sequence can be obtained by attaching ears serially or by bar-amalgamation of a cactus to $G_N$. The resulting graph is denoted $G_{r,s,t}$. Finally we prove that the Euler-genus distribution of $G_{r,s,t}$ is asymptotically normal distribution.
A bar-amalgamation $G \oplus_e H$ of two disjoint graphs $H$ and $G$ is obtained by adding a new edge $e = uv$ between a vertex $u$ of $G$ and a vertex $v$ of $H$. 

Figure: $K_4 \oplus_e (K_5 - 2K_2)$
Let \( \{H_n\}_{n=1}^{\infty} \) be a sequence of connected graphs. A sequence of tree-like graphs \( \{G_n\}_{n=1}^{\infty} \) are obtained in the following ways:

(A) For \( n = 1 \), \( G_1 = H_1 \).

(B) If we have obtained the graph \( G_{n-1} \), the graph \( G_n \) is obtained by adding an edge between a vertex of \( G_{n-1} \) and a vertex of \( H_n \). I.e, \( G_n = G_{n-1} \oplus e H_n \).
Proposition 1 Let $\{\xi_k\}_{k=1}^{\infty}$ be a sequence of independent random variables with finite second moments. Let $
abla_k = E\xi_k^2 - (E\xi_k)^2 > 0, B_n = \sqrt{\sum_{k=1}^{n} \sigma_k^2}$. Assume that for a sequence of positive constants $\{C_n\}_{n=1}^{\infty}$, we have $\sup_{1 \leq k \leq n} |\xi_k| \leq C_n$, and $\lim_{n \to \infty} \frac{C_n}{B_n} = 0$. Then, it holds that

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \frac{\sum_{k=1}^{n} (\xi_k - E\xi_k)}{B_n} \leq x \right) - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right| = 0$$

when $n$ tends to infinity.
Theorem (A)

Let \( \{G_n\}_{n=1}^{\infty} \) be a sequence of tree-like graphs. Assume the followings holds.

- For each \( n \), \( H_n \) is a finite connected graph.
- Except a finite number of \( n \in N \), \( \gamma_{\text{max}}(H_n) > \gamma_{\text{min}}(H_n) \).

Then, the genus distribution of \( G_n \) is asymptotically normal distribution with mean \( e_n \) and variance \( v_n \).
Sketch of proof. Step 1. In this step, we prove that: for two positive constants \( c, C \), it holds that

\[
0 < c \leq \inf_{n \in \mathbb{N}} \Gamma_{\text{var}}(H_n) \leq \sup_{n \in \mathbb{N}} \Gamma_{\text{var}}(H_n) \leq C \quad (3)
\]

and

\[
0 < c \leq \inf_{n \in \mathbb{N}} \Gamma_{\text{avg}}(H_n) \leq \sup_{n \in \mathbb{N}} \Gamma_{\text{avg}}(H_n) \leq C. \quad (4)
\]
Step 2. For each $n \in N$, the distribution of random variable $\xi_1 + \cdots + \xi_n$ is given by

$$P(\xi_1 + \cdots + \xi_n = j) = \frac{\gamma_j(G_n)}{\Gamma_{G_n}(1)}, j = 0, 1, \cdots .$$

Step 3. In this step, we give a proof of our theorem using Proposition 1.
Thank you!