The Fréchet Mean of Inhomogeneous Random Graphs

François G. Meyer(✉)

Applied Mathematics, University of Colorado Boulder, Boulder, CO 80305, USA
fmeyer@colorado.edu
https://francoismeyer.github.io

Abstract. To characterize the "average" of a set of graphs, one can compute the sample Fréchet mean. We prove the following result: if we use the Hamming distance to compute distances between graphs, then the Fréchet mean of an ensemble of inhomogeneous random graphs is obtained by thresholding the expected adjacency matrix: an edge exists between the vertices i and j in the Fréchet mean graph if and only if the corresponding entry of the expected adjacency matrix is greater than $1/2$. We prove that the result also holds for the sample Fréchet mean when the expected adjacency matrix is replaced with the sample mean adjacency matrix. This novel theoretical result has some significant practical consequences; for instance, the Fréchet mean of an ensemble of sparse inhomogeneous random graphs is the empty graph.

Keywords: Fréchet mean · Statistical network analysis

1 Introduction

The Fréchet mean graph has become a standard tool for the analysis of graph-valued data (e.g., [5,6,8,10,13,14]). In this work, we derive the expression for the population Fréchet mean for inhomogeneous Erdős-Rényi random graphs [2]. We prove that the sample Fréchet mean is consistent, and could be estimated using a simple thresholding rule. This novel theoretical result implies that the sample Fréchet mean computed from a training set of graphs, which display specific topological features of interest, will not inherit from the training set the desired topological structure.

We consider the set $\mathcal{G}$ formed by all undirected unweighted simple labeled graphs with vertex set $\{1, \ldots, n\}$. We denote by $\mathcal{S}$ the set of $n \times n$ adjacency matrices of graphs in $\mathcal{G}$,

$$\mathcal{S} = \left\{ \boldsymbol{A} \in \{0,1\}^{n \times n}; \text{where } a_{ij} = a_{ji}, \text{and } a_{i,i} = 0;\ 1 \leq i < j \leq n \right\}. \quad (1)$$

F. G. Meyer—Supported by the National Science Foundation (CCF/CIF1815971).

R. M. Benito et al. (Eds.): COMPLEX NETWORKS 2021, SCI 1015, pp. 207–219, 2022.
https://doi.org/10.1007/978-3-030-93409-5_18

We denote by $\mathcal{G}(n, \boldsymbol{P})$ the probability space formed by the inhomogeneous Erdős-Rényi random graphs [2], defined on $\{1, \ldots, n\}$, where a graph G with adjacency matrix $\boldsymbol{A}$ has probability,

$$\mathbb{P}(\boldsymbol{A}) = \prod_{1 \leq 1 < j \leq n} [p_{ij}]^{a_{ij}} [1 - p_{ij}]^{1 - a_{ij}} . \tag{2}$$

The $n \times n$ matrix $\boldsymbol{P} = [p_{ij}]$ determines the edge probabilities $0 \leq p_{ij} \leq 1$, with $p_{ii} = 0$. We identify $\mathcal{G}(n, \boldsymbol{P})$ with the probability space $(\mathcal{S}, \mathbb{P})$, where $\mathbb{P}$ is defined by (2). The prominence of $\mathcal{G}(n, \boldsymbol{P})$ stems from its ability to provide tractable models of random graphs that can capture many of the structures of real networks (e.g., stochastic block models, which have great practical importance). We equip $\mathcal{G}$ with the Hamming distance defined as follows.

Definition 1. *The Hamming distance between G and G' in $\mathcal{G}$, with adjacency matrix $\boldsymbol{A}$ and $\boldsymbol{A}'$ respectively, is given by*

$$d_H(G, G') = \sum_{1 \leq i < j \leq n} |a_{ij} - a'_{ij}|. \tag{3}$$

We characterize the mean of the probability $\mathbb{P}$ with the Fréchet mean graph, [7].

Definition 2. *The Fréchet mean of the probability measure $\mathbb{P}$ is the set formed by the solutions to*

$$\boldsymbol{\mu}[\mathbb{P}] = \operatorname*{argmin}_{G \in \mathcal{G}} \sum_{G' \in \mathcal{G}} d_H^2(G, G') \mathbb{P}(G') \tag{4}$$

where d_H is the Hamming distance (1).

By replacing $\mathbb{P}$ with the empirical measure, the concept of Fréchet mean graph can be extended to a sample of graphs defined on the same vertex set $\{1, \ldots, n\}$.

Definition 3. *Let $\left\{G^{(k)}\right\}_{1 \leq k \leq N}$, be independent random graphs, sampled from $\mathbb{P}$. The sample Fréchet mean is the set composed of the solutions to*

$$\widehat{\boldsymbol{\mu}}_N[\mathbb{P}] = \operatorname*{argmin}_{G \in \mathcal{G}} \frac{1}{N} \sum_{k=1}^{N} d^2(G, G^{(k)}). \tag{5}$$

We note that solutions to the minimization problems (4) and (5) always exist, but need not be unique. Because all the results in this paper hold for any graph in the set formed by the solutions to (4) and (5), and without any loss of generality, we assume that $\boldsymbol{\mu}[\mathbb{P}]$ and $\widehat{\boldsymbol{\mu}}_N[\mathbb{P}]$ each contains a single element.

This notion of centrality is well adapted to metric spaces (e.g., [4,10,13]). The vital role played by the Fréchet mean as a location parameter, is exemplified

in the works of [1,14], who have created novel families of random graphs by generating random perturbations around a given Fréchet mean.

Because the focus of this work is not the computation of the Fréchet mean graph, but rather a theoretical analysis of the properties that the Fréchet mean graph inherits from the probability measure $\mathbb{P}$, defined in (2), we can assume that all the graphs are defined on the same vertex set.

1.1 Our Main Contributions

The prominence of the inhomogeneous Erdős-Rényi random graph model [2] prompts the following critical question: does the Fréchet mean of $\mathbb{P}$ inherit from the probability space $\mathcal{G}(n, \boldsymbol{P})$ any of the edge connectivity information encoded by $\boldsymbol{P}$?

In this paper, we answer this question. We show in Theorem 1 that the population Fréchet mean graph $\boldsymbol{\mu}[\mathbb{P}]$ can be obtained by thresholding the mean adjacency matrix $\mathbb{E}[\boldsymbol{A}] = \boldsymbol{P}$; an edge exists between the vertices i and j in $\boldsymbol{\mu}[\boldsymbol{A}]$ if and only if $\mathbb{E}[\boldsymbol{A}]_{ij} > 1/2$. We prove in Theorem 2 that this result also holds for the sample Fréchet mean graph, $\widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}]$, when $\mathbb{E}[\boldsymbol{A}]$ is replaced with the sample mean adjacency matrix, $\widehat{\mathbb{E}}_N[\boldsymbol{A}]$.

2 Main Results

Let $\boldsymbol{P} = [p_{ij}]$ be an $n \times n$ symmetric matrix with entries $0 \leq p_{ij} \leq 1$. In the following two theorems we determine the Fréchet mean graph, and sample mean graph, of graphs in $\mathcal{G}(n, \boldsymbol{P})$. In the following, we denote by $[n]$ the set $\{1, \ldots, n\}$.

2.1 The Population Fréchet Mean Graph of $\mathcal{G}(n, \mathbf{P})$

Theorem 1. *The Fréchet mean graph $\boldsymbol{\mu}[\mathbb{P}]$ of the probability measure (2), is given by*

$$\forall i, j \in [n], \quad \boldsymbol{\mu}[\mathbb{P}]_{ij} = \begin{cases} 1 & \text{if} \quad \mathbb{E}[\boldsymbol{A}]_{ij} = p_{ij} > 1/2, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Proof. The proof is given in Sect. 3.2.

2.2 The Sample Fréchet Mean Graph of a Graph Sample in $\mathcal{G}(n, \mathbf{P})$

We now turn our attention to the sample Fréchet mean graph, which has recently been used for the statistical analysis of graph-valued data (e.g., [5,8,14,17]). The computation of the sample Fréchet mean graph using the Hamming distance is NP-hard [3]. For this reason, several alternatives have been proposed (e.g., [6,8]).

Before presenting the second result, we take a short detour through the sample Fréchet median of the probability measure $\mathbb{P}$, [9,11,16], minimiser of

$$\widehat{\boldsymbol{m}}_N[\mathbb{P}] = \underset{G \in \mathcal{G}}{\operatorname{argmin}} \frac{1}{N} \sum_{k=1}^{N} d_H(G, G^{(k)}), \tag{7}$$

and which can be computed using the majority rule [1].

Lemma 1. *The adjacency matrix* $\widehat{\boldsymbol{m}}_N[\boldsymbol{A}]$ *of* $\widehat{\boldsymbol{m}}_N[\mathbb{P}]$ *is given by*

$$\forall i,j \in [n], \quad \widehat{\boldsymbol{m}}_N[\boldsymbol{A}]_{ij} = \begin{cases} 1 & \text{if } \sum_{k=1}^N a_{ij}^{(k)} \geq N/2, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

We now come back to the second main contribution, where we prove that the sample Fréchet mean graph of N independent random graphs sampled from $\mathcal{G}(n,\boldsymbol{P})$ is asymptotically equal (for large sample size) to the sample Fréchet median graph, with high probability.

Theorem 2. $\forall \delta \in (0,1), \exists N_\delta, \forall N \geq N_\delta$, $\widehat{\boldsymbol{m}}_N[\boldsymbol{A}]$ *and* $\widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}]$ *are given by*

$$\forall i,j \in [n], \quad \widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}]_{ij} = \widehat{\boldsymbol{m}}_N[\boldsymbol{A}]_{ij} = \begin{cases} 1 & \text{if } \mathbb{E}[\boldsymbol{A}]_{ij} = p_{ij} > 1/2, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

with probability $1-\delta$ *over the realizations of the graphs,* $\{G^{(1)},\ldots,G^{(N)}\}$ *in* $\mathcal{G}(n,\boldsymbol{P})$.

Proof. The proof is given in Sect. 3.5.

The practical impact of Theorem 2 is given by the following corollary, which is an elementary consequence of Theorem 2 and Lemma 1.

Corollary 1. $\forall \delta \in (0,1), \exists N_\delta, \forall N \geq N_\delta$, $\widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}]$ *is given by the majority rule,*

$$\forall i,j \in [n], \quad \widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}]_{ij} = \begin{cases} 1 & \text{if } \sum_{k=1}^N a_{ij}^{(k)} > N/2, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

with probability $1-\delta$ *over the realizations of the graphs,* $\{G^{(1)},\ldots,G^{(N)}\}$, *in* $\mathcal{G}(n,\boldsymbol{P})$.

3 Proofs of the Main Results

We give in the following the proofs of Theorems 1 and 2. In the process, we prove several technical lemmata.

3.1 The Population and sample Fréchet Functions

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two adjacency matrices in $\mathcal{S}$. We provide below an expression for the Hamming distance squared, $d_H^2(\boldsymbol{A},\boldsymbol{B})$, where the computation is split between the entries of $\boldsymbol{A}$ along the edges of $\boldsymbol{B}$, $\mathcal{E}(\boldsymbol{B})$, and the entries of $\boldsymbol{A}$ along the "nonedges" of $\boldsymbol{B}$, $\overline{\mathcal{E}}(\boldsymbol{B})$. We denote by $|\mathcal{E}(\boldsymbol{B})|$ the number of edges in $\boldsymbol{B}$.

Lemma 2. *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ two matrices in $\mathcal{S}$. Then,*

$$d_H^2(\boldsymbol{A},\boldsymbol{B}) = \Big[\sum_{1\le i<j\le n} a_{ij}\Big]^2 + |\mathcal{E}(\boldsymbol{B})|^2 + 2\,|\mathcal{E}(\boldsymbol{B})|\left[\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} a_{ij} - \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} a_{ij}\right] - 4\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}\ \sum_{(i',j')\in\overline{\mathcal{E}}(\boldsymbol{B})} a_{ij}a_{i'j'} \tag{11}$$

We now define the (population) Fréchet function associated with (4).

Definition 4. *We denote by F_2 the Fréchet function associated with the Fréchet mean*

$$F_2(\boldsymbol{B}) = \sum_{\boldsymbol{A}\in\mathcal{S}} d_H^2(\boldsymbol{A},\boldsymbol{B})\mathbb{P}(\boldsymbol{A})\,. \tag{12}$$

As explained in the following lemma, the value of the Fréchet function $F_2(\boldsymbol{B})$ depends only on the entries of the probability matrix $\boldsymbol{P}$ along the edges of $\boldsymbol{B}$.

Lemma 3. *Let $\boldsymbol{B}\in\mathcal{S}$, let $\mathcal{E}(\boldsymbol{B})$ be the set of edges of the graph associated to $\boldsymbol{B}$. Then*

$$F_2(\boldsymbol{B}) = \Big[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}(1-2p_{ij}) + \sum_{1\le i<j\le n} p_{ij}\Big]^2 + \sum_{1\le i<j\le n} p_{ij}(1-p_{ij}). \tag{13}$$

Proof. The proof of (13) relies on the expression for the Hamming distance squared, (11). The proof is omitted; instead we will prove Lemma 4, the proof of which is extremely similar, albeit more technical, to that of Lemma 3.

3.2 Proof of Theorem 1

We are now in position to prove the first theorem. By Lemma 3, we seek the matrix $\boldsymbol{B}$, with edge set $\mathcal{E}(\boldsymbol{B})$, that minimizes the Fréchet function defined by (13). Let us denote

$$x \stackrel{\text{def}}{=} \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}(1-2p_{ij})\,. \tag{14}$$

Since $0\le p_{ij}\le 1$, x is confined to the following interval,

$$-\sum_{1\le i<j\le n} p_{ij} \le -\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} p_{ij} \le x \le \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} 1 \le n(n-1)/2. \tag{15}$$

In fact, $x = -\sum_{1\le i<j\le n} p_{ij}$, only if $\forall i,j\in[n],\ p_{ij}=1$, and the graph associated to $\boldsymbol{B}$ is the complete graph. This case is of no interest to us, and thus we can assume that $\boldsymbol{P}$ is always chosen such that

$$-\sum_{1\le i<j\le n} p_{ij} < x. \tag{16}$$

We define,

$$f(x) \stackrel{\text{def}}{=} \Big[x + \sum_{1 \le i < j \le n} p_{ij}\Big]^2.$$

We have

$$f(x) = F_2(\boldsymbol{B}) - \sum_{1 \le i < j \le n} p_{ij}(1 - p_{ij}).$$

Minimizing F_2 is therefore equivalent to minimizing f. Clearly, $f(x)$ is convex, has a global minimum at $x_{\min} = -\sum_{1 \le i < j \le n} p_{ij}$, and is increasing for $x \ge -\sum_{1 \le i < j \le n} p_{ij}$. We seek x^* that minimizes $f(x)$ over the interval wherein x is enclosed,

$$\Big(-\sum_{1 \le i < j \le n} p_{ij},\ n(n-1)/2\Big].$$

We note that because of (16), $x_{\min} < x^*$. Also, x^* cannot be positive; otherwise, we would get $f(x^*) > f(0)$. The optimal value x^* is obtained by minimizing the distance from x^* to $-\sum_{1 \le i < j \le n} p_{ij}$,

$$x^* - (-\sum_{1 \le i < j \le n} p_{ij}) = \sum_{(i,j) \in \mathcal{E}(\boldsymbol{B})} (1 - 2p_{ij}) + \sum_{1 \le i < j \le n} p_{ij} \ge \sum_{(i,j);1-2p_{ij}<0} (1 - 2p_{ij}) + \sum_{1 \le i < j \le n} p_{ij}. \tag{17}$$

The lower bound (17) is independent of $\boldsymbol{B}$, and can be attained by choosing,

$$\boldsymbol{\mu}\big[\mathbb{P}\big]_{ij} = \begin{cases} 1 & \text{if} \quad p_{ij} > 1/2, \\ 0 & \text{otherwise}, \end{cases} \tag{18}$$

as advertised in the theorem. □

3.3 The Sample Fréchet Function for the Hamming Distance

We now consider N independent random graphs, $\{G^{(k)}\}_{1 \le k \le N}$, sampled from $\mathcal{G}(n, \boldsymbol{P})$, with adjacency matrices $\boldsymbol{A}^{(k)}$. The sample Fréchet function $\widehat{F}_2(\boldsymbol{B})$ associated with the sample Fréchet mean graph is defined as follows.

Definition 5. *We denote by $\widehat{F}_2$ the sample Fréchet function*

$$\widehat{F}_2(\boldsymbol{B}) = \frac{1}{N} \sum_{k=1}^{N} d_H^2(\boldsymbol{A}^{(k)}, \boldsymbol{B}). \tag{19}$$

We have the following expression for $\widehat{F}_2(\boldsymbol{B})$, which is similar to (13).

Lemma 4. *Let $\boldsymbol{B} \in \mathcal{S}$, let $\mathcal{E}(\boldsymbol{B})$ be the set of edges of the graph associated to $\boldsymbol{B}$. Then*

$$\widehat{F}_2(\boldsymbol{B}) = \left[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \left[1 - 2\widehat{\mathbb{E}}_N\left[a_{ij}\right]\right] + \sum_{1\le i<j\le n} \widehat{\mathbb{E}}_N\left[a_{ij}\right] \right]^2 + \sum_{1\le i<j\le n} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\left(1 - \widehat{\mathbb{E}}_N\left[a_{ij}\right]\right) \tag{20}$$

$$- \sum_{1\le i<j\le n} \sum_{\substack{1\le i'<j'\le n \\ (i,j)\neq(i',j')}} \left(\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right)$$

$$+ 4 \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \sum_{(i',j')\in\overline{\mathcal{E}}(\boldsymbol{B})} \left(\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right) \tag{21}$$

where the sample mean and sample correlation are defined by

$$\widehat{\mathbb{E}}_N\left[a_{ij}\right] = \frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} \quad \textit{and} \quad \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right] = \frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} a_{i'j'}^{(k)} \tag{22}$$

Proof. The proof is similar to the proof of Lemma 3. For each graph $G^{(k)}$, we apply Eq. (11), sum over all the graphs in the sample, and divide by N,

$$\widehat{F}_2(\boldsymbol{B}) = |\mathcal{E}(\boldsymbol{B})|^2 + 2\,|\mathcal{E}(\boldsymbol{B})| \left[\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} \frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} - \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} \right]$$
$$+ \frac{1}{N}\sum_{k=1}^{N}\left[\sum_{1\le i<j\le n} a_{ij}^{(k)} \right]^2 - 4 \sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} \sum_{(i',j')\in\mathcal{E}(\boldsymbol{B})} \left[\frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} a_{i'j'}^{(k)} \right]. \tag{23}$$

Using the expressions for the sample mean and correlation, in (22), we get

$$\widehat{F}_2(\boldsymbol{B}) = |\mathcal{E}(\boldsymbol{B})|^2 + 2\,|\mathcal{E}(\boldsymbol{B})| \left[\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right] - \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right] \right]$$
$$+ \frac{1}{N}\sum_{k=1}^{N}\left[\sum_{1\le i<j\le n} a_{ij}^{(k)} \right]^2 - 4 \sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} \sum_{(i',j')\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right] \tag{24}$$

We note that

$$\frac{1}{N}\sum_{k=1}^{N}\left[\sum_{1\le i<j\le n} a_{ij}^{(k)} \right]^2 = \sum_{1\le i<j\le n} \sum_{1\le i'<j'\le n} \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]. \tag{25}$$

Also, we have

$$
\begin{aligned}
&|\mathcal{E}(\boldsymbol{B})|^2 + 2\,|\mathcal{E}(\boldsymbol{B})|\left[\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right] - \sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\right] \\
&\quad = \left[|\mathcal{E}(\boldsymbol{B})| - 2\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\right]^2 - \sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] \\
&\qquad + 4\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})}\ \sum_{(i',j')\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] \qquad (26)
\end{aligned}
$$

We can then substitute (25) and (26) into (24), and we get

$$
\begin{aligned}
\widehat{F}_2(\boldsymbol{B}) = &\left[|\mathcal{E}(\boldsymbol{B})| - 2\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\right]^2 \\
&- \sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n}\left[\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right] \\
&+ 4\sum_{(i,j)\in\overline{\mathcal{E}}(\boldsymbol{B})}\ \sum_{(i',j')\in\mathcal{E}(\boldsymbol{B})}\left[\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right] \qquad (27)
\end{aligned}
$$

Finally, we can extract from the second line of (27), the term that corresponds to $(i,j) = (i',j')$, and we get

$$
\begin{aligned}
&\sum_{1\le i<j\le n}\ \sum_{1\le i'<j'\le n}\left[\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right] \\
&\qquad = \sum_{1\le i<j\le n}\ \sum_{\substack{1\le i'<j'\le n \\ (i',j')\neq(i,j)}} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right] \\
&\qquad\quad + \sum_{1\le i<j\le n} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\left(\widehat{\mathbb{E}}_N\left[a_{ij}\right] - 1\right). \qquad (28)
\end{aligned}
$$

Substituting (28) into the second line of (27), we obtain (21) as announced in Lemma 4. □

3.4 Concentration of the Sample Fréchet Function

In the following lemma, we show that for large sample size N, the sample Fréchet function $\widehat{F}_2(\boldsymbol{B})$ concentrates around its population counterpart, $F_2(\boldsymbol{B})$.

Lemma 5. *$\forall\delta\in(0,1), \exists N_\delta, \forall N\ge N_\delta, \forall \boldsymbol{B}\in\mathcal{S}$,*

$$
\widehat{F}_2(\boldsymbol{B}) = \left[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}(1-2p_{ij}) + \sum_{1\le i<j\le n} p_{ij}\right]^2 + \sum_{1\le i<j\le n} p_{ij}\,(1-p_{ij}) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \qquad (29)
$$

with probability $1-\delta$ over the realization of the sample $\left\{G^{(k)}\right\}_{1\le k\le N}$.

Proof. The sample mean $\widehat{\mathbb{E}}_N[a_{ij}]$, defined in (22), is the sum of Bernoulli random variables, and it concentrates around its mean p_{ij}. We use Hoeffding inequality to bound the variation of $\widehat{\mathbb{E}}_N[a_{ij}]$ around p_{ij}. For each $1 \leq i < j \leq n$, we have,

$$\mathbb{P}\left(\boldsymbol{A}^{(k)} \sim \mathcal{G}(n, \boldsymbol{P}); \left|\widehat{\mathbb{E}}_N[a_{ij}] - p_{ij}\right| \geq \varepsilon\right) \leq \exp\left(-2N\varepsilon^2\right). \tag{30}$$

To control $\sum_{k=1}^{N} a_{ij}^{(k)}$ for all $1 \leq i < j < n$, we use a union bound and we get,

$$\forall 1 \leq i < j < n, \quad \left|\widehat{\mathbb{E}}_N[a_{ij}] - p_{ij}\right| \leq \frac{\alpha}{\sqrt{N}}, \tag{31}$$

with probability $1 - \delta/8$, and where $\alpha = \sqrt{\log(2n/\sqrt{\delta})}$. We now study the concentration of the sample correlation,

$$\widehat{\mathbb{E}}_N[\rho_{ij,i'j'}] = \frac{1}{N}\sum_{k=1}^{N} a_{ij}^{(k)} a_{i'j'}^{(k)}, \tag{32}$$

when the pair of edges (i,j) and (i',j') are distinct. Because $(i,j) \neq (i',j')$, the terms $a_{ij}^{(k)}$ and $a_{i'j'}^{(k)}$ are always independent, and the product $a_{ij}^{(k)} a_{i'j'}^{(k)}$ is a Bernoulli random variable with parameter $p_{ij}p_{i'j'}$. We conclude that $\widehat{\mathbb{E}}_N[\rho_{ij,i'j'}]$ is the sum of Bernoulli random variables, and concentrates around its mean.

We use Hoeffding inequality to bound the variation of $\widehat{\mathbb{E}}_N[\rho_{ij,i'j'}]$. Replicating the argument used for $\widehat{\mathbb{E}}_N[a_{ij}]$ mutatis mutandis, yields

$$\forall\, 1 \leq i < j \leq n, \forall\, 1 \leq i' < j' \leq n, \ \left|\widehat{\mathbb{E}}_N[\rho_{ij,i'j'}] - p_{ij}p_{i'j'}\right| \leq \frac{\beta}{\sqrt{N}}, \tag{33}$$

with probability $1 - \delta/8$, where $\beta = \sqrt{\log(n^2/\sqrt{\delta/2})}$. In summary, we have

$$\begin{aligned} &\forall\, 1 \leq i < j \leq n, \quad \forall\, 1 \leq i' < j' \leq n, \quad \text{with} \quad (i,j) \neq (i',j'), \\ &\widehat{\mathbb{E}}_N[a_{ij}] = p_{ij} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \text{ and } \widehat{\mathbb{E}}_N[\rho_{ij,i'j'}] = p_{i'j'} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \end{aligned} \tag{34}$$

with probability $1 - \delta/4$. We are now in position to substitute $\widehat{\mathbb{E}}_N[a_{ij}]$ and $\widehat{\mathbb{E}}_N[\rho_{ij,i'j'}]$ with the expressions given by (34), in $\widehat{F}_2(\boldsymbol{B})$ given by (21) in Lemma 4. Using (34), the first term in (21) becomes

$$\begin{aligned} &\Big[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}\left(1 - 2\widehat{\mathbb{E}}_N[a_{ij}]\right) + \sum_{1\leq i<j\leq n}\widehat{\mathbb{E}}_N[a_{ij}]\Big]^2 \\ &\qquad = \Big[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}(1 - 2p_{ij}) + \sum_{1\leq i<j\leq n} p_{ij}\Big]^2 + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \end{aligned} \tag{35}$$

with probability $1-\delta/4$. Also, we have

$$\sum_{1\leq i<j\leq n} \widehat{\mathbb{E}}_N\left[a_{ij}\right]\left(1-\widehat{\mathbb{E}}_N\left[a_{ij}\right]\right) = \sum_{1\leq i<j\leq n} p_{ij}\left(1-p_{ij}\right) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \tag{36}$$

with probability $1-\delta/4$. The last two terms in (21) can be neglected since,

$$\begin{aligned}\sum_{1\leq i<j\leq n} \sum_{\substack{1\leq i'<j'\leq n\\(i,j)\neq(i',j')}} &\left[\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right] \\ &= \sum_{1\leq i<j\leq n} \sum_{\substack{1\leq i'<j'\leq n\\(i,j)\neq(i',j')}} \left[p_{ij}p_{i'j'} - p_{ij}p_{i'j'}\right] + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),\end{aligned} \tag{37}$$

with probability $1-\delta/4$. Similarly

$$\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})} \sum_{(i',j')\in\overline{\mathcal{E}}(\boldsymbol{B})} \left[\widehat{\mathbb{E}}_N\left[a_{ij}\right]\widehat{\mathbb{E}}_N\left[a_{i'j'}\right] - \widehat{\mathbb{E}}_N\left[\rho_{ij,i'j'}\right]\right] = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right), \tag{38}$$

with probability $1-\delta/4$. Substituting (35), (36), (37), and (38) into (21) yields the following estimate

$$\widehat{F}_2(\boldsymbol{B}) = \Big[\sum_{(i,j)\in\mathcal{E}(\boldsymbol{B})}(1-2p_{ij}) + \sum_{1\leq i<j\leq n} p_{ij}\Big]^2 + \sum_{1\leq i<j\leq n} p_{ij}\left(1-p_{ij}\right) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),$$

which holds with probability $1-\delta$. □

3.5 Proof of Theorem 2

We prove (9), in Theorem 2, for the sample Fréchet mean. The proof for the sample Fréchet median is completely similar (it also uses a concentration of measure argument for the Fréchet function defined in (7)) and is therefore omitted.

Because of Lemma 5, (29) implies that

$$\forall\delta\in(0,1), \exists N_\delta, \forall N\geq N_\delta, \forall \boldsymbol{B}\in\mathcal{S}, \widehat{F}_2(\boldsymbol{B}) = F_2(\boldsymbol{B}) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),$$

with probability $1-\delta$ over the realization of the sample $\left\{G^{(k)}\right\}_{1\leq k\leq N}$. For N large enough, the main term dominates the expression of $\widehat{F}_2(\boldsymbol{B})$, and we can neglect the $\mathcal{O}\left(1/\sqrt{N}\right)$ term. We are left with $F_2(\boldsymbol{B})$, the Fréchet function for the population mean, given by (13), in Lemma 3. The minimum of $\widehat{F}_2(\boldsymbol{B})$ is thus achieved for the adjacency matrix given by the population Fréchet mean, $\boldsymbol{\mu}[\mathbb{P}]$, defined by (6), as advertised in (9), in Theorem 2. □

4 Simulation Studies

We compare our theoretical analysis to finite sample estimates, which were computed using numerical simulations. The software used to conduct the experiments is publicly available [15].

All graphs were generated using the $\mathcal{G}(n, \boldsymbol{P})$ model (2). We varied the edge probability matrix, $\boldsymbol{P}$. For each simulation, $\boldsymbol{P}$ was chosen randomly using independent (up to symmetry) beta random variables, $p_{ij} \sim \text{beta}(2, 10)$. The sample Fréchet mean was computed using the approximation provided by (10). All graphs had $n = 512$ vertices. We varied the sample size for $N \in [10, 7079]$. For each sample size N, we first generated a probability matrix $\boldsymbol{P}$ from the beta distribution, and we then sampled N independent random graphs $G^{(1)}, \ldots, G^{(N)}$ from $\mathcal{G}(n, \boldsymbol{P})$.

We illustrate the concentration of the sample Fréchet function for large N, described by Lemma 5. Figure 1 displays the mean error between the population Fréchet function $F_2(\boldsymbol{B})$, given by (13), and the sample Fréchet function $\widehat{F}_2(\boldsymbol{B})$, given by (21), as a function of the sample size N. The average error between $F_2(\boldsymbol{B})$ and $\widehat{F}_2(\boldsymbol{B})$, is computed using a sample of $N_B = 16$ independent random graphs $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_{N_B}$, sampled from $\mathcal{G}(n, \boldsymbol{P})$,

$$\widehat{\mathbb{E}}_{N_B}\big[F_2(\boldsymbol{B}) - \widehat{F}_2(\boldsymbol{B})\big] = \frac{1}{N_B}\sum_{i=1}^{N_B}\Big|F_2(\boldsymbol{B}_i) - \widehat{F}_2(\boldsymbol{B})\Big|. \tag{39}$$

For each N, the sample average error $\widehat{\mathbb{E}}_{N_B}\big[F_2(\boldsymbol{B}) - \widehat{F}_2(\boldsymbol{B})\big]$, corresponds to a point in Fig. 1-left. We repeated this simulation 64 times to create 64 different values of the error (39). A linear regression was computed and is displayed (in blue) in Fig. 1-left. The slope of the error was found to be -0.5028, confirming the $1/\sqrt{N}$ decay of the error predicted by Lemma 5.

To estimate the deviation of the sample Fréchet mean graph away from the population Fréchet mean graph, we computed the Hamming distance between the population Fréchet mean graph (6), and the sample Fréchet mean graph (9). Figure 1-right displays $d_H(\boldsymbol{\mu}[\boldsymbol{A}], \widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}])$. A linear regression was computed and is displayed (in blue) in Fig. 1-right. The slope was found to be -0.6707, suggesting that the sample Fréchet mean converges toward the population Fréchet mean at a rate $1/N^{2/3}$, which is faster than the rate predicted by Lemma 5.

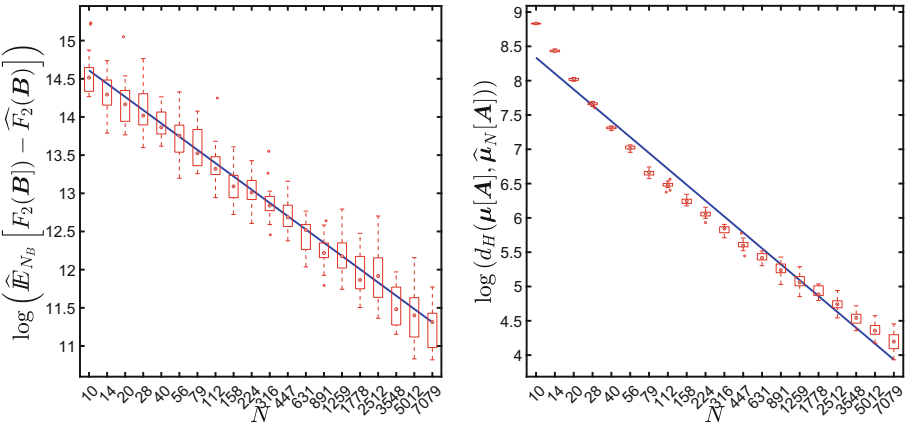

Fig. 1. Left: mean error $\widehat{\mathbb{E}}_{N_B}\left[F_2(\boldsymbol{B}) - \widehat{F}_2(\boldsymbol{B})\right]$ between the population and the sample Fréchet functions, as a function of the sample size N. Right: $d_H(\boldsymbol{\mu}[\boldsymbol{A}], \widehat{\boldsymbol{\mu}}_N[\boldsymbol{A}])$, the Hamming distance between the population Fréchet mean graph and the sample Fréchet mean graph.

5 Discussion and Conclusion

Our answer to the question of the authors in [14]: "what is the "mean" network (rather than how do we estimate the success-probabilities of an inhomogeneous random graph), and do we want the "mean" itself to be a network?" is therefore disappointing in the context of the probability space $\mathcal{G}(n, \boldsymbol{P})$. While the Fréchet mean is indeed an element of $\mathcal{G}(n, \boldsymbol{P})$, it only provides a simplistic sketch of that probability space. Consider for instance sparse graphs where $\min p_{ij} < 1/2$ (e.g., graphs with $o(n^2)$ but $\omega(n)$ edges), then the sample Fréchet mean is the empty graph, and is pointless.

On a more positive note, our analysis provides a theoretical justification for several algorithms designed to recover a graph from noisy measurements of its adjacency matrix. For instance, the authors in [12] compute the sample mean of the noisy adjacency matrices, and threshold the sample mean to recover an unweighted graph. Our results offer a theoretical justification of the approach of [12]: Theorem 2 proves that the algorithm described in [12] recovers the sample Fréchet mean graph.

References

1. Banks, D., Constantine, G.: Metric models for random graphs. J. Classif. **15**(2), 199–223 (1998)
2. Bollobás, B., Janson, S., Riordan, O.: The phase transition in inhomogeneous random graphs. Random Struct. Algorithms **31**(1), 3–122 (2007)
3. Chen, J., Hermelin, D., Sorge, M.: On computing centroids according to the p-norms of Hamming distance vectors. In: 27th Annual European Symposium on Algorithms (ESA 2019), vol. 144, pp. 28:1–28:16, Dagstuhl, Germany (2019)

4. Chowdhury, S., Mémoli, F.: The metric space of networks (2018)
5. Dubey, P., Müller, H.G.: Fréchet change-point detection. Ann. Stat. **48**(6), 3312–3335 (2020)
6. Ferrer, M., Valveny, E., Serratosa, F., Riesen, K., Bunke, H.: Generalized median graph computation by means of graph embedding in vector spaces. Pattern Recogn. **43**(4), 1642–1655 (2010)
7. Fréchet, M.: Les espaces abstraits et leur utilité en statistique théorique et même en statistique appliquée. Journal de la Société Française de Statistique **88**, 410–421 (1947)
8. Ginestet, C.E., Li, J., Balachandran, P., Rosenberg, S., Kolaczyk, E.D.: Hypothesis testing for network data in functional neuroimaging. Ann. Appl. Stat. **11**(2), 725–750 (2017)
9. Han, F., Han, X., Liu, H., Caffo, B., et al.: Sparse median graphs estimation in a high-dimensional semiparametric model. Ann. Appl. Stat. **10**(3), 1397–1426 (2016)
10. Jain, B.J.: Statistical graph space analysis. Pattern Recogn. **60**, 802–812 (2016)
11. Jiang, X., Munger, A., Bunke, H.: On median graphs: properties, algorithms, and applications. IEEE Trans. Pattern Anal. Mach. Intell. **23**(10), 1144–1151 (2001)
12. Josephs, N., Li, W., Kolaczyk, E.D.: Network recovery from unlabeled noisy samples (2021)
13. Kolaczyk, E.D., Lin, L., Rosenberg, S., Walters, J., Xu, J., et al.: Averages of unlabeled networks: geometric characterization and asymptotic behavior. Ann. Stat. **48**(1), 514–538 (2020)
14. Lunagómez, S., Olhede, S.C., Wolfe, P.J.: Modeling network populations via graph distances. J. Am. Stat. Assoc. 1–18 (2020). Published online: 08 Sep 2020. https://www.tandfonline.com/doi/full/10.1080/01621459.2020.1763803
15. Meyer, F.G.: The Mean of Inhomogeneous Random Graphs (2021). https://github.com/francoismeyer/frechet-mean
16. Mukherjee, L., Singh, V., Peng, J., Xu, J., Zeitz, M.J., Berezney, R.: Generalized median graphs and applications. J. Comb. Optim. **17**(1), 21–44 (2009)
17. Zambon, D., Alippi, C., Livi, L.: Change-point methods on a sequence of graphs. IEEE Trans. Signal Process. **67**(24), 6327–6341 (2019)