

MAX-PLUS OBJECTS TO STUDY THE COMPLEXITY OF GRAPHS

CRISTIANO BOCCI*, LUCA CHIANTINI†, AND FABIO RAPALLO‡

Abstract. Given an undirected graph G , we define a new object H_G , called the mp-chart of G , in the max-plus algebra. We use it, together with the max-plus permanent, to describe the complexity of graphs. We show how to compute the mean and the variance of H_G in terms of the adjacency matrix of G and we give a central limit theorem for H_G . Finally, we show that the mp-chart is easily tractable also for the complement graph.

Key words. permanent of adjacency matrices; combinatorial central limit theorem; random permutations; complement graph.

AMS subject classifications. 05C30, 60F05.

1. Introduction. The work presented in this paper has been inspired by the need of simple and actual techniques to measure the complexity of a graph, especially in the case of sparse graphs. This problem arises in several fields of applications, from Computer Science to Economics, from Biology to Social Sciences. As general references for the graph theory, we mention in particular the books [7], [8], [5] and [10], where the reader can find the main mathematical achievements in the theory. For a general survey on recent applications of graph theory see, for instance, [1]. The reader interested in some more technical papers can refer to [16] for applications to Economics, to [9] for applications to Econophysics, to [17] for applications to Biology, and to [15] for applications to Molecular Biology. In such papers, sparse graphs play a prominent role.

Although an undirected graph $G = (V, E)$ is a rather simple structure, consisting of a set V of N vertices and a set of edges $E \subset V \times V$, in graph theory there are several different approaches, depending on the specific application we are looking for. In particular, a graph can be fixed or random, depending on whether the elements in E are random or not. Moreover, many efforts have been done to analyze dynamical graphs, where the vertex set V and/or the edge set E vary with time, see e.g. [11].

Here, we restrict our analysis to fixed graphs. In this framework, there are interesting developments in the area of Combinatorics, about the study of the properties of $0-1$ matrices. These matrices naturally arise in the framework of graphs, as the adjacency matrix A_G of the fixed graph G . Some recent developments in this direction, with applications to graph theory, are described in [2], [3], and [4].

In the present paper, we investigate some questions about undirected graph, in order to study and describe their structure, with special attention to sparse graphs. Our work is related to the matching problem. As a preliminary remark, we argue that, for sparse graphs, the classical descriptors of the complexity, such as the degree distribution and the permanent of the adjacency matrix do not give actual information. Thus, we use the max-plus arithmetic and the corresponding expression of the permanent, and we show that this object is more suitable for describing of the complexity for sparse graph. The use of the max-plus arithmetic naturally leads to the definition

*Dipartimento di Scienze Matematiche e Informatiche “R. Magari”, Università di Siena, Pian dei Mantellini 44, 53100 Siena, Italy (cristiano.bocci@unisi.it)

†Dipartimento di Scienze Matematiche e Informatiche “R. Magari”, Università di Siena, Pian dei Mantellini 44, 53100 Siena, Italy (luca.chiantini@unisi.it)

‡Dipartimento di Scienze e Tecnologie Avanzate, Università del Piemonte Orientale, Viale Teresa Michel 11, 15121 Alessandria, Italy (fabio.rapallo@mfn.unipmn.it)

of a more complete index of the structure of a graph, and we define a vector called the *mp-chart of the graph*. This vector is nothing else but a probability distribution, and we show that it converges to a Normal distribution through a combinatorial Central Limit Theorem. Several examples on small- and medium-sized graphs are given to show that our definitions are easy to apply and provide practical information about the complexity of the structure of the graph under study. All the computations have been carried out with Maple, see [18], and R, see [19]. All the simulations come from simple R routines, without any additional package.

This paper is only concerned with undirected fixed graphs. Nevertheless, the same strategy can be applied to other situations, such as bipartite graphs, undirected unfixed graphs, random graphs, and so on.

The paper is organized as follows. In Section 2 we define the max-plus permanent of an adjacency matrix (i.e., the permanent under the max-plus arithmetic), we state its main properties, we study its connections with the classical permanent, and we discuss some simple examples. In Section 3 we define a new object associated to a graph, and we call it the mp-chart of the graph. We compute its mean and variance, and we show that, under suitable conditions, it converges to a Normal distribution as the size of the graph goes to infinity. Some simple simulations show that the convergence is quite good also for small values of the size. In Section 4, we show how the mp-chart of a graph is related to the mp-chart of the complement graph. Finally, Section 5 is devoted to suggest some future directions of this research.

2. The max-plus permanent. Let $G = (V, E)$ be an undirected graph with N vertices. Let A_G be the $N \times N$ adjacency matrix of G , defined by $(A_G)_{i,j} = 1$ if $(i, j) \in E$ and 0 otherwise. In the classic definition of undirected graph, the matrix A_G is symmetric and with zero diagonal entries, as we do not consider loops.

As mentioned in the Introduction, the study of the complexity of a given graph is one of the most relevant problems about graphs in Applied Probability. This analysis can be performed through the distribution of the degrees (i.e., the number of edges involving each vertex) and through the permanent (or the determinant) of the adjacency matrix A_G .

The determinant of A_G is

$$\det(A_G) = \sum_{\pi} (-1)^{|\pi|} \prod_{i=1}^N (A_G)_{i, \pi(i)},$$

where the sum is taken over all the permutations π of $\{1, \dots, N\}$ and $|\pi|$ denotes the parity of π . The permanent of A_G is

$$\text{perm}(A_G) = \sum_{\pi} \prod_{i=1}^N (A_G)_{i, \pi(i)}.$$

The use of permanent to describe the complexity of a graph is justified by the following well-known property.

PROPOSITION 2.1. *The permanent of A_G is the number of bijections $\phi : V \rightarrow V$ compatible with E , i.e. such that $(v, \phi(v)) \in E$ for all $v \in V$.*

In fact, $\text{perm}(A_G)$ is the number of permutations π with $A_{1, \pi(1)} = \dots = A_{N, \pi(N)} = 1$ and the permutation π is just the bijection ϕ in the proposition.

However, the analysis based on the degree distribution and the permanent is not adequate for sparse graphs. In fact, it is enough to have an isolated vertex to produce

a null permanent. Nevertheless, it is interesting to study the structure of a sparse graph.

To overcome this difficulty, we make use of the tropicalization of the permanent. In the classical settings, Tropical Arithmetic is defined through the operations:

$$x \oplus y = \min\{x, y\} \quad x \otimes y = x + y$$

But, with Tropical Arithmetic, the determinant (or permanent) of an adjacency matrix is always 0, because of the nullity of the main diagonal of A_G .

Thus, we use the max-plus algebra, with operations:

$$x \oplus y = \max\{x, y\} \quad x \otimes y = x + y$$

Consequently, the explicit expression of the max-plus permanent is

$$\text{perm}_{\text{mp}}(A_G) := \bigoplus_{\pi} \left(\bigotimes_{i=1}^N (A_G)_{i, \pi(i)} \right) = \max_{\pi} \sum_{i=1}^N (A_G)_{i, \pi(i)} \quad (2.1)$$

The max-plus permanent is the maximum over $N!$ terms. Each of the $N!$ terms is the sum of N terms in $\{0, 1\}$. Thus, the max-plus permanent $\text{perm}_{\text{mp}}(A_G)$ is zero if and only if the matrix A_G is the null matrix. On the opposite side, the maximum allowed value of the max-plus permanent is N .

The use of the max-plus permanent to analyze sparse graphs has a first reason in the following property.

LEMMA 2.2. *The following relation holds:*

$$\text{perm}_{\text{mp}}(A_G) = N \iff \text{perm}(A_G) > 0. \quad (2.2)$$

Proof. $\text{perm}_{\text{mp}}(A_G) = N$ if and only if there exists a permutation π such that $\sum_{i=1}^N (A_G)_{i, \pi(i)} = N$. This happens if and only if there exists π such that $(A_G)_{i, \pi(i)} = 1$ for $i = 1, \dots, N$, i.e. if and only if $\text{perm}(A_G) > 0$. \square

REMARK 2.3. *Notice that, from Lemma 2.2 and from the previous discussion, it follows that the max-plus permanent is able to discriminate among graphs with standard permanent equal to zero.*

Moreover, we explicitly write the following consistency property, whose simple proof is straightforward.

LEMMA 2.4. *Let G and H be two graphs on two disjoint sets of vertices. Then,*

$$\text{perm}_{\text{mp}}(A_{G \cup H}) = \text{perm}_{\text{mp}}(A_G) + \text{perm}_{\text{mp}}(A_H). \quad (2.3)$$

The max-plus permanent has interesting connections with the subgraphs. Let $G' = (V', E')$ a graph. If $V' \subset V$ and $E' \subset E$, then G' is a subgraph of $G = (V, E)$. A subgraph $G' = (V', E')$ is the subgraph induced by V' if E' contains all the edges in E involving the vertices in V' . In order to analyze the max-plus permanent, in view of Equation (2.1), we introduce here the notion of t -term, which is strictly related to the subgraphs of G . Such connections will be studied later in this section.

DEFINITION 2.5. *Given a graph G with adjacency matrix A_G , a t -term is a sequence of indices $(i_1, j_1) \cdots (i_t, j_t)$ with*

- $1 \leq i_1 < \dots < i_t \leq N$;

- the j_k 's, with $1 \leq j_k \leq N$ are all distinct;
- $(A_G)_{i_k, j_k} = 1$ for all k .

For a t -term P , we denote $I(P) = \{i_1, \dots, i_t\}$ and $J(P) = \{j_1, \dots, j_t\}$.

Roughly speaking, a t -term corresponds to a sequence of positions of t ones in the permutations. A straightforward consequence is the following statement.

PROPOSITION 2.6. *The max-plus permanent of G is q if and only if there exists a q -term and there are no t -term with $t > q$.*

PROPOSITION 2.7. *Let t the maximum integer such that there exists a t -term, then there exists a t -term P such that $I(P) = J(P)$.*

Proof. We prove the statement by induction on t . If $t = 2$ there is nothing to prove since if the 2-term is given by $(i_1, j_1)(i_2, j_2)$ it is enough to consider the 2-term $(i_1, j_1)(j_1, i_1)$.

Consider now a t -term Q and suppose there exists a k such that i_k is in $I(Q) \setminus J(Q)$. First of all we notice that j_k must be in $I(Q)$. If not, we can add the element (j_k, i_k) to Q obtaining a $(t+1)$ -term which is a contradiction, since Q is maximal. Hence, since $j_k \in I(Q)$ then there exists a s such that $j_k = i_s \in I(Q)$. Then, we substitute (i_s, j_s) with (j_k, i_k) in our t -term and we obtain a new t -term of the form $(i_k, j_k)(j_k, i_k)Q'$ where Q' is a $(t-2)$ -term. This term Q' arises from the sub-matrix A' of $A(G)$ where we remove rows and columns i_k and j_k . Hence Q' is a maximal $(t-2)$ -term for A' . By induction, the proof follows. \square

REMARK 2.8. *If P is a 3-term, then we must have $I(P) = J(P)$. In fact, if P is $(i_1, j_1)(i_2, j_2)(i_3, j_3)$ with $I(P) \neq J(P)$ then, by the previous proposition, we obtain a new 3-term $(i_k, j_k)(j_k, i_k)(i_3, j_3)$. Then it would be possible to extend it to $(i_k, j_k), (j_k, i_k)(i_3, j_3)(j_3, i_3)$ against the maximality of P .*

REMARK 2.9. *In view of Proposition 2.7, the max-plus permanent is the cardinality of the largest subset of V with a bijection compatible with E . This is another way to see that the max-plus permanent is able to detect the complexity of the graphs with null classical permanent.*

Denote by ℓ_G the number of edges of a graph G . Among the subgraphs of Proposition 2.7, we are mainly interested in the ones with a minimal number of edges. These subgraphs are maximal in term of $\text{perm}_{\text{mp}}(A_{G'})$, but minimal in term of $\ell_{G'}$. We made this more precise by the following definition.

DEFINITION 2.10. *An mp-maximal subgraph G' of a graph G , is a subgraph of G with $q = \text{perm}_{\text{mp}}(A_G)$ vertices,*

$$\text{perm}_{\text{mp}}(A_{G'}) = \text{perm}_{\text{mp}}(A_G) \quad (2.4)$$

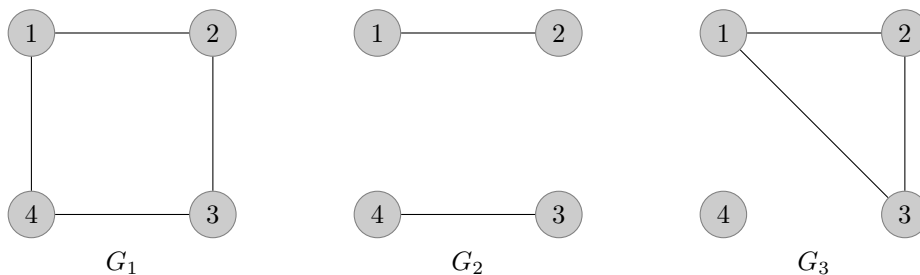
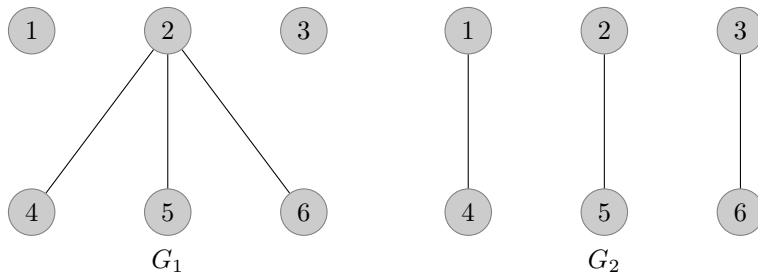
and for all other subgraph G'' of G satisfying (2.4) one has $\ell_{G'} \leq \ell_{G''}$.

The rest of this section is devoted to the discussion of some examples and some useful remarks. In order to understand the definitions introduced above, we start with some small graphs.

EXAMPLE 2.11. *Let us analyze the three graphs on 4 vertices drawn in Figure 2.1. Their adjacency matrices are respectively*

$$A_{G_1} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \quad A_{G_2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad A_{G_3} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In the first graph, all vertices are connected and $\text{perm}_{\text{mp}}(A_{G_1}) = 4$. However, this is not the minimal way to obtain a max-plus permanent equal to 4. In fact, it is


 FIG. 2.1. The graphs G_1 , G_2 and G_3 for Example 2.11.

 FIG. 2.2. The two graphs G_1 and G_2 for Example 2.12.

easy to check that $\text{perm}_{\text{mp}}(A_{G_2}) = 4$. Thus the graph G_2 represents a mp-maximal subgraph for the graph G_1 , but it is not the only one. If we look now at the graph G_3 , we notice that there is a cycle of length 3 and an isolated vertex. In such case, we have $\text{perm}_{\text{mp}}(A_{G_3}) = 3$, and there is only one mp-maximal subgraph.

EXAMPLE 2.12. To illustrate the behavior of the max-plus permanent and of the mp-maximal subgraphs, we analyze two opposite examples. with the same length. The two graphs are drawn in Figure 2.2. The graph G_1 on the left is the union of a tree and two isolated vertices, with $\text{perm}_{\text{mp}}(A_{G_1}) = 2$ and 4 maximal subgraphs with two vertices and one edge each. On the opposite side, the graph G_2 has a perfect matching, $\text{perm}_{\text{mp}}(A_{G_2}) = 6$ and there is only 1 maximal subgraph, i.e., the graph G_2 itself.

PROPOSITION 2.13. Two mp-maximal subgraphs are not disjoint.

Proof. Consider a graph G such that $\text{perm}_{\text{mp}}(A_G) = t$ and let $G' = (V', E')$ and $G'' = (V'', E'')$ be two mp-maximal subgraphs of G with t vertices each. Suppose that V' and V'' are disjoint. Then, by Formula (2.3), the adjacency matrix of $G' \cup G''$ has max-plus permanent $2t$. Then $\text{perm}_{\text{mp}}(A_G) \geq 2t$ which is a contradiction. \square

REMARK 2.14. In the max-plus arithmetic, the definition of determinant is not unique, see [13]. Therefore, one has to define the positive and negative determinant. In particular, the positive max-plus determinant is the maximum of the sums $\sum_i A_{i, \pi(i)}$ over all even permutations π . The negative determinant is defined by taking the odd permutations instead of the even ones. This issue is another reason to use the permanent instead of the determinant in the max-plus environment.

3. The mp-chart of a graph. The information about a graph is not contained only in the max-plus permanent, but in the whole distribution of the $N!$ terms $\sum_{i=1}^N (A_G)_{i, \pi(i)}$. Thus, in this section we define the mp-chart of a graph as the distribution of the $N!$ terms above, and we prove that this distribution converges to a Gaussian distribution through the Hoeffding's combinatorial central limit theorem,

see [14]. We also show that the mean and the variance of that distribution can be computed easily from the adjacency matrix.

DEFINITION 3.1. *Let G be a graph and A_G its adjacency matrix. Let $h_G(k)$ be the number of permutations π such that $\sum_{i=1}^N (A_G)_{i,\pi(i)} = k$. We call the $(N+1)$ -dimensional integer vector $H_G = (h_G(0), \dots, h_G(N))$ the mp-chart of the graph G .*

This object captures many features of the graph and has some relevant theoretical properties. To understand the meaning of H_G , notice that $h_G(k)$ is just the number of permutations π such that the sequence $(A_G)_{1,\pi(1)}, \dots, (A_G)_{N,\pi(N)}$ contains a k -term but not a $(k+1)$ -term. This gives precisely the meaning and the usefulness of the notion of random permutation in that context.

EXAMPLE 3.2. *We use here a very simple scheme inspired by Econophysics, see [21] and [12]. Consider a population with N agents, each possessing one good. The goods can be sent and received only along the edges of a graph G and each agent can possess only one good. Given a random permutation π of $\{1, \dots, N\}$, the i -th agent can send its good to $\pi(i)$ if it receive a good from $\pi^{-1}(i)$. The quantity $\sum_{i=1}^N (A_G)_{i,\pi(i)}$ is exactly the number of agents involved in this process. Of course, similar examples can be adapted to many other sciences.*

We start the analysis of the mp-chart with the study of the mean $\mathbb{E}(H_G)$ and the variance $\mathbb{V}(H_G)$. Although these computations could be carried out applying Theorem 2 in [14], it is useful to state explicitly the proof for adjacency matrices.

Given an adjacency matrix $A = (A_G)$, note that, for a random permutation π , the object

$$S_N(\pi) = A_{1,\pi(1)} + \dots + A_{N,\pi(N)} \quad (3.1)$$

is the sum of N binary random variables, and $A_{i,\pi(i)}$ is the addendum chosen in the i -th row of the adjacency matrix.

In order to analyze the mean of H_G , defined by

$$\mathbb{E}(H_G) = \frac{\sum_{k=0}^N kh_G(k)}{N!}, \quad (3.2)$$

it is convenient to adopt an inductive approach.

THEOREM 3.3. *The mean of the mp-chart H_G is*

$$\mathbb{E}(H_G) = \frac{2\ell_G}{N}, \quad (3.3)$$

where ℓ_G is the number of edges in the graph G .

Proof. Clearly, if A_G is the null matrix, then $\mathbb{E}(H_G) = 0$. Suppose that the formula (3.3) holds true for $\ell_G - 1$. By direct inspection, adding one edge has the following consequences. Among the $N!$ terms $S_N(\pi)$:

- $(N-2)!$ of them increase by 2;
- $2(N-2)(N-2)!$ of them increase by 1;
- the remaining $((N-2)^2 + N-1)(N-2)!$ do not change.

Thus,

$$\mathbb{E}(H_G) = \frac{2(\ell_G - 1)}{N} + \frac{2(N-2)! + 2(N-2)(N-2)!}{N!} = \frac{2(\ell_G - 1)}{N} + \frac{2}{N} = \frac{2\ell_G}{N}.$$

□

$$\left(\begin{array}{cc|cc} 0 & 0 & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots \\ \hline \vdots & \vdots & & \\ \vdots & \vdots & & \end{array} \right) \quad \left(\begin{array}{cc|cc} 0 & 1 & \cdots & \cdots \\ 1 & 0 & \cdots & \cdots \\ \hline \vdots & \vdots & & \\ \vdots & \vdots & & \end{array} \right)$$

Case (a) Case (b)

FIG. 3.1. The two cases arising in the proof of Theorem 3.5.

EXAMPLE 3.4. Given a complete graph G , its adjacency matrix has 0 on the diagonal and 1 elsewhere. The graph has $N(N-1)/2$ edges. Hence $\mathbb{E}(H_G) = N-1$ which is the maximum allowed.

Notice that the mean $\mathbb{E}(H_G)$ depends only in the number of edges of G , whatever they are collocated, that is, $\mathbb{E}(H_G)$ does not take into account the topology of the graph. On the other hand, the variance $\mathbb{V}(H_G)$ depends on the position of the edges.

THEOREM 3.5. The variance of the mp-chart H_G is

$$\mathbb{V}(H_G) = \sum_{i=1}^N \frac{d_i(N-d_i)}{N^2} + \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{d_i d_j - N T_{i,j}}{N^2(N-1)}, \quad (3.4)$$

where d_1, \dots, d_N are the degrees of the vertices and $T_{i,j} = \langle r_i, r_j \rangle$ is the scalar product of the i -th and the j -th row of A_G .

Proof. By direct computation, the formula (3.4) holds for $N \leq 2$.

To prove the validity of Eq. (3.4) for $N \geq 3$, it is enough to compute the covariances

$$\begin{aligned} \text{Cov}(A_{i,\pi(i)}, A_{j,\pi(j)}) &= \mathbb{E}(A_{i,\pi(i)} A_{j,\pi(j)}) - d_i d_j / N^2 = \\ &= \mathbb{P}(A_{i,\pi(i)} = 1, A_{j,\pi(j)} = 1) - d_i d_j / N^2. \end{aligned}$$

Without loss of generality we can fix $i = 1$ and $j = 2$ and we write for brevity A_1 for $A_{1,\pi(1)}$ and A_2 for $A_{2,\pi(2)}$. Moreover, we suppose that d_1 and d_2 are both non zero. (If $d_1 = 0$ or $d_2 = 0$, then trivially $\text{Cov}(A_1, A_2) = 0$).

We divide the computation in two cases, and to help the reader we have sketched the two cases in Figure 3.1.

- Case (a): $(1, 2)$ is not an edge of the graph. Then:
 - there are $2(N-1)!$ permutations such that $\pi(1) = 1$ or $\pi(1) = 2$. For all these cases, $(A_1 = 1, A_2 = 1)$ is impossible;
 - there are $2(N-2)(N-2)!$ permutations such that $\pi(1) > 2$, but $\pi(2) = 1$ or $\pi(2) = 2$. Also in all these cases, $(A_1 = 1, A_2 = 1)$ is impossible;
 - there are $(N-2)(N-3)(N-2)!$ permutations such that $\pi(1) > 2$ and $\pi(2) > 2$, and among these permutations

$$(T_{1,2} d_2 + (d_1 - T_{1,2}) d_2)(N-2)! = (d_1 d_2 - T_{1,2})(N-2)!$$

are such that $A_1 A_2 = 1$.

Therefore,

$$\text{Cov}(A_1, A_2) = \frac{d_1 d_2 - NT_{1,2}}{N^2(N-1)}$$

- Case (b): $(1, 2)$ is an edge of the graph. Then:
 - there are $(N-1)!$ permutations such that $\pi(1) = 1$. In all such cases, $(A_1 = 1, A_2 = 1)$ is impossible;
 - there are $(N-2)!$ permutations such that $\pi(1) = 2$ and $\pi(2) = 1$. For such permutations, $A_1 = 1$ and $A_2 = 1$;
 - there are $(N-2)(N-2)!$ permutations such that $\pi(1) = 2$ and $\pi(2) > 2$. Among these permutations, $(d_2 - 1)(N-2)!$ are such that $A_1 A_2 = 1$.
 - there are $(N-2)(N-2)!$ permutations such that $\pi(1) > 2$ and $\pi(2) = 1$. Among these permutations, $(d_1 - 1)(N-2)!$ are such that $A_1 A_2 = 1$.
 - there are $(N-2)(N-2)!$ permutations such that $\pi(1) > 2$ and $\pi(2) = 2$. In all such cases, $A_1 A_2 = 0$.
 - there are $(N-2)(N-3)(N-2)!$ permutations such that $\pi(1) > 2$ and $\pi(2) > 2$, and among these

$$\begin{aligned} & (T_{1,2}(d_2 - 2) + (d_1 - 1 - T_{1,2})(d_2 - 1))(N-2)! = \\ & = (d_1 d_2 - d_1 - d_2 - T_{1,2} + 1)(N-2)! \end{aligned}$$

are such that $A_1 A_2 = 1$.

Therefore, adding up all the contributions, we obtain again

$$\text{Cov}(A_1, A_2) = \frac{d_1 d_2 - NT_{1,2}}{N^2(N-1)}$$

The formula in Eq. (3.4) is now straightforward. \square

EXAMPLE 3.6. Consider the matrices

$$A_{G_1} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad A_{G_2} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The graph G_1 has two consecutive edges, while the graph G_2 has two disjoint edges. An easy computation gives

$$H_{G_1} = (48, 48, 24, 0, 0, 0)$$

and

$$H_{G_2} = (53, 44, 18, 4, 1, 0)$$

with equal means $\mathbb{E}(H_{G_1}) = \mathbb{E}(H_{G_2}) = 4/5$. On the contrary, the variances are $\mathbb{V}(H_{G_1}) = 14/25$ and $\mathbb{V}(H_{G_2}) = 19/25$, respectively.

EXAMPLE 3.7. As a second example, consider the two graphs on the set vertices $V = \{1, \dots, 7\}$ shown in Figure 3.2. Notice that G_1 and G_2 differ by only one edge. The two mp-charts H_{G_1} and H_{G_2} have the same mean and variance, namely

$$\mathbb{E}(H_{G_1}) = \mathbb{E}(H_{G_2}) = \frac{12}{7} \quad \mathbb{V}(H_{G_1}) = \mathbb{V}(H_{G_2}) = \frac{170}{147}$$

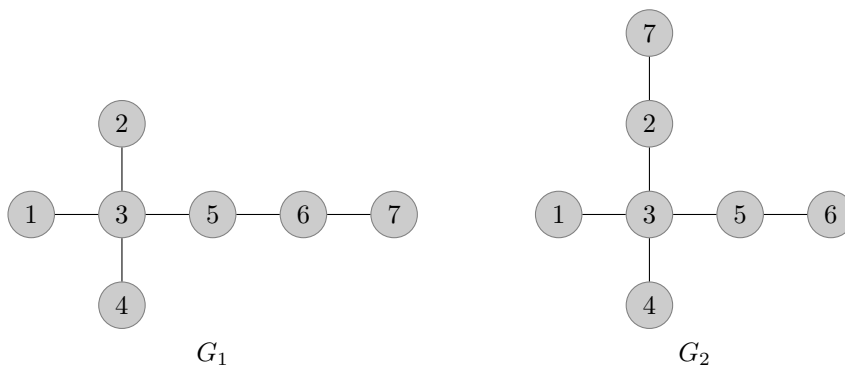


FIG. 3.2. Two graphs with different mp-charts, but with equal means and variances.

but the mp-charts are different:

$$\begin{aligned}
 H_{G_1} &= (678, 1512, 1716, 840, 294, 0, 0, 0) \\
 H_{G_2} &= (674, 1480, 1792, 840, 218, 32, 4, 0).
 \end{aligned}$$

The results above lead to a central limit theorem.

THEOREM 3.8. *Let G_N be a graph with N vertices and let A_{G_N} be its adjacency matrix. Let π be a random permutation of $\{1, \dots, N\}$ chosen with uniform probability and define*

$$S_N(\pi) = \sum_{i=1}^N (A_{G_N})_{i, \pi(i)}. \quad (3.5)$$

If $\mathbb{V}(S_N)$ goes to infinity as $N \rightarrow \infty$, then the distribution of S_N is asymptotically normal.

Proof. We make use of Theorem 3 in [14]. Define the auxiliary matrix R with elements

$$R_{i,j} = (A_{G_N})_{i,j} - \frac{d_i}{N} - \frac{d_j}{N} + \frac{1}{N^2} \sum_{h,k} (A_{G_N})_{h,k} \quad (3.6)$$

Then, a sufficient condition for the asymptotic normality is that

$$\lim_{N \rightarrow \infty} \frac{\max_{1 \leq i, j \leq N} R_{i,j}^2}{\frac{1}{N} \sum_{i,j=1}^N R_{i,j}^2} = 0. \quad (3.7)$$

Now observe that the numerator is bounded, as $-2 \leq R_{i,j} \leq 2$ for all i and j . Moreover, Theorem 2 in the same paper [14] states that

$$\mathbb{V}(S_N) = \frac{1}{N-1} \sum_{i,j=1}^N R_{i,j}^2. \quad (3.8)$$

Combining these facts, the result follows. \square

REMARK 3.9. *Note that in our problem one can not use the classical central limit theorems based on α -mixing sequences or m -dependent variables, see for instance [6,*

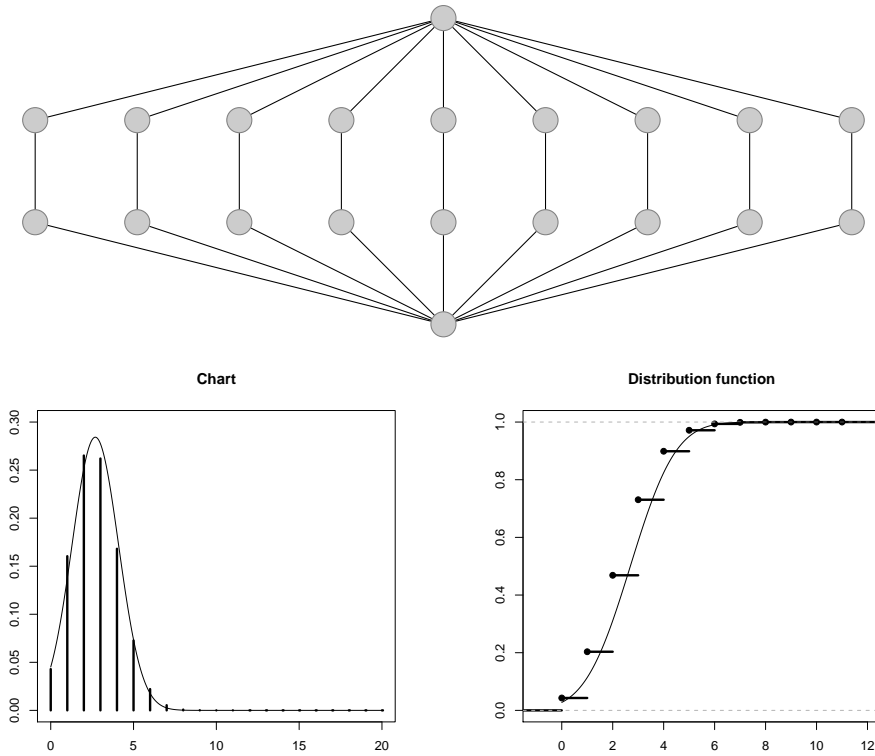


FIG. 3.3. Checking the convergence for $N = 20$: First graph and its results.

Ch. 27] and [20]. Indeed, the covariance between $A_{i,\pi(i)}$ and $A_{j,\pi(j)}$ does not vanish as $|i - j|$ goes to infinity.

In order to inspect the behavior of the convergence to the Gaussian distribution, we have computed the mp-chart for some graphs with 20 vertices.

The three examples in Figures 3.3-3.5 show that the convergence is quite good, meaning that the Gaussian approximation is valid also for medium-sized graphs.

Two remarks are needed to understand the examples: (a) The mp-chart is approximated through a standard Monte Carlo technique, sampling 100,000 random permutations. This number is considerably smaller than the total number of permutations ($20! \cong 10^{18}$), but it provides quite accurate approximations; (b) The results are presented through two plots, showing the mp-chart (normalized to 1) and its distribution function, both compared with the appropriate Normal distribution.

The first graph corresponds to an adjacency matrix with block structure. The graph and the two plots of the results are presented in Figure 3.3.

The second graph has a different shape, as it corresponds to an adjacency matrix with band structure. The graph and the two plots of the results are presented in Figure 3.4.

For the third graph analyzed here, we present only the results. The graph has been constructed with $N(N - 1)/4 = 95$ edges randomly chosen among the 190 edges of the complete graph with uniform probability. The results are shown in Figure 3.5.

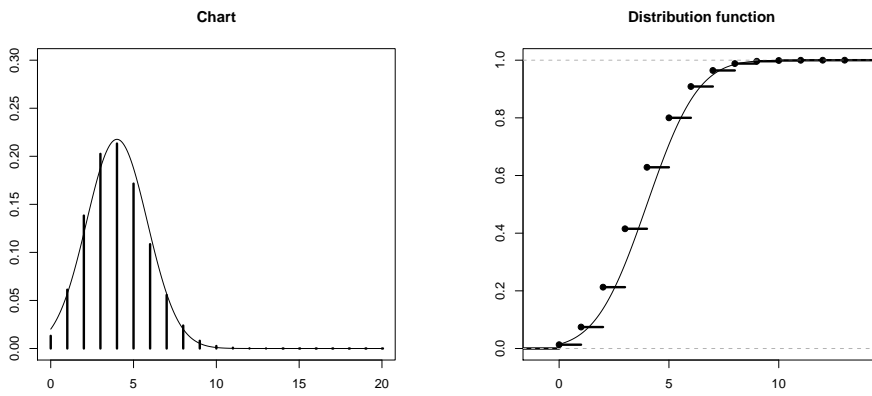
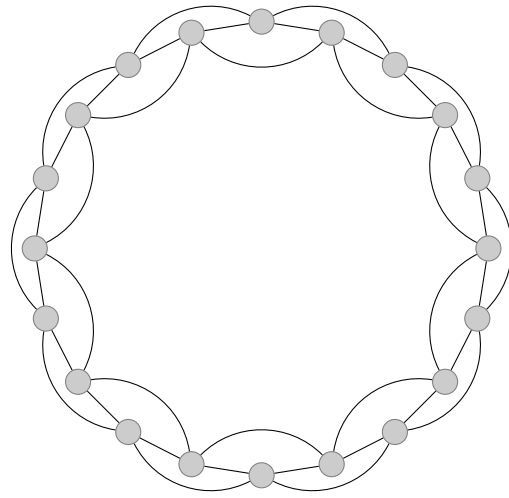


FIG. 3.4. Checking the convergence for $N = 20$: Second graph and its results.

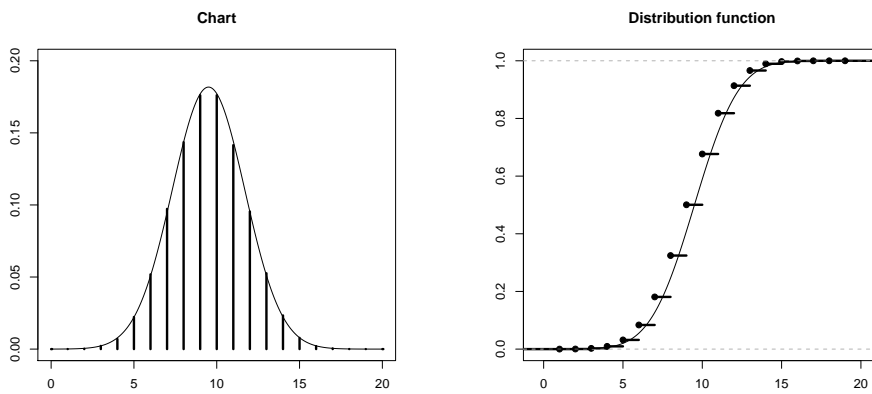


FIG. 3.5. Checking the convergence for $N = 20$: The results for the third graph.

4. The mp-chart of the complement graph. In the literature, the complement of a graph $G = (V, E)$ is a graph on the same vertex set V and the set of edges $V^2 \setminus E$. Since our starting graph has no loop (e, e) , this forces G^c to contain all of them. To avoid this problem we give a different definition of complement graph, more useful for our purposes.

DEFINITION 4.1. *Given a graph $G = (V, E)$, its complement graph G^c is a graph on the same vertex set V and the set of edges $(V^2 \setminus \Delta) \setminus E$, where $\Delta \subset E^2$ is the diagonal set, i.e. $\Delta = \{(v, v) : v \in V\}$.*

REMARK 4.2. *From the previous definition, we notice that, if G and G^c have respectively ℓ_G and ℓ_{G^c} edges, then $\ell_G + \ell_{G^c} = N(N - 1)/2$.*

As mentioned in Section 3, there are some nice properties linking the mp-chart H_G of a graph G with the mp-chart of its complement G^c . To study these connections, we start with a preliminary lemma.

LEMMA 4.3. *Let G be a graph with N vertices. Denote by T_{ij}^c the scalar product between the i -th row and j -th column in the adjacency matrix of G^c . The following formula relates the quantities T_{ij}^c and T_{ij} :*

$$T_{ij}^c = T_{ij} + N - 2 - d_i - d_j + \langle r_i, E_j \rangle + \langle r_j, E_i \rangle. \quad (4.1)$$

where the E_i 's are the vectors in the canonical basis of \mathbb{R}^N .

Proof. Define the vector

$$v_{ij} = r_i - r_j - \langle r_i - r_j, E_i \rangle E_i - \langle r_i - r_j, E_j \rangle E_j \quad (4.2)$$

Since $\langle r_i - r_j, E_i \rangle = \langle r_i, E_i \rangle - \langle r_j, E_i \rangle$ and $\langle r_t, E_t \rangle = 0$ (in fact, the t -th coordinate of r_t is zero, while E_t has a 1 in the t -th coordinate and 0 elsewhere), we can write

$$v_{ij} = r_i - r_j + \langle r_j, E_i \rangle E_i - \langle r_i, E_j \rangle E_j$$

The scalar product $\langle v_{ij}, v_{ij} \rangle$ measures the number of positions, out of the diagonal, where r_i and r_j are different. Thus, if we denote by T_{ij}^c the scalar product of the corresponding lines, r_i^c, r_j^c in the complement graph, one has

$$T_{ij}^c = N - 2 - T_{ij} - \langle v_{ij}, v_{ij} \rangle. \quad (4.3)$$

Now, we substitute in the previous formula the expression of v_{ij} given in (4.2), and we obtain:

$$\begin{aligned} T_{ij}^c &= N - 2 - T_{ij} - \langle v_{ij}, v_{ij} \rangle = N - 2 - T_{ij} - \\ &\quad - \langle r_i - r_j + \langle r_j, E_i \rangle E_i - \langle r_i, E_j \rangle E_j, r_i - r_j + \langle r_j, E_i \rangle E_i - \langle r_i, E_j \rangle E_j \rangle. \end{aligned}$$

Noting that $\langle r_t, r_t \rangle = d_t$ and $\langle r_i, r_j \rangle = \langle r_j, r_i \rangle = T_{ij}$, a straightforward computation leads to

$$T_{ij}^c = T_{ij} + N - 2 - d_i - d_j + \langle r_i, E_j \rangle^2 + \langle r_j, E_i \rangle^2.$$

Since, for all i, j , with $i \neq j$, the value of $\langle r_i, E_j \rangle$ can be either 0 or 1, we can remove the squares from the previous formula, leading to Equation (4.1). \square

THEOREM 4.4. *Given a graph G with N vertices,*

- (a) $\mathbb{E}(H_{G^c}) = N - 1 - \mathbb{E}(H_G)$;
- (b) $\mathbb{V}(H_{G^c}) = \mathbb{V}(H_G) + 1 - 2\mathbb{E}(H_G)/(N - 1)$.

Proof. To prove part (a), it is enough to use Theorem 3.3 and Remark 4.2. One has

$$\mathbb{E}(H_{G^c}) = \frac{2\ell_{G^c}}{N} = \frac{N(N-1) - 2\ell_G}{N} = N - 1 - \mathbb{E}(H_G).$$

To prove part (b), we apply Theorem 3.5 to the graph G^c . Therefore we have:

$$\mathbb{V}(H_{G^c}) = \sum_{i=1}^N \frac{d_i^c(N - d_i^c)}{N^2} + \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{d_i^c d_j^c - NT_{i,j}^c}{N^2(N-1)}.$$

The degree d_i^c of a vertex in the complement graph is given by $d_i^c = N - 1 - d_i$. Using also Lemma 4.3 one can write $\mathbb{V}(H_{G^c})$ as

$$\begin{aligned} \mathbb{V}(H_{G^c}) &= \sum_{i=1}^N \frac{(N-1-d_i)(d_i+1)}{N^2} + \\ &+ \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{(N-1-d_i)(N-1-d_j) - N(T_{ij} + N-2-d_i-d_j) + \langle r_i, E_j \rangle + \langle r_j, E_i \rangle}{N^2(N-1)} \end{aligned} \quad (4.4)$$

The first sum in formula (4.4) can be written as

$$\sum_{i=1}^N \frac{(d_i)(N-d_i)}{N^2} + \sum_{i=1}^N \frac{(N-1-2d_i)}{N^2},$$

where an easy computation shows that

$$\sum_{i=1}^N \frac{(N-1-2d_i)}{N^2} = \frac{N-1}{N} - \frac{4\ell_G}{N^2}.$$

About the second sum in formula (4.4), we observe that each term can be expressed as

$$\begin{aligned} &\frac{(N-1-d_i)(N-1-d_j) - N(T_{ij} + N-2-d_i-d_j) + \langle r_i, E_j \rangle + \langle r_j, E_i \rangle}{N^2(N-1)} = \\ &= \frac{d_i d_j - NT_{i,j} + d_i + d_j + 1 - N(\langle r_i, E_j \rangle + \langle r_j, E_i \rangle)}{N^2(N-1)}. \end{aligned}$$

Since

$$\sum_{\substack{i,j=1 \\ i \neq j}}^N \langle r_i, E_j \rangle + \langle r_j, E_i \rangle = 4\ell_G \quad \text{and} \quad \sum_{\substack{i,j=1 \\ i \neq j}}^N (d_i + d_j) = 4\ell_G(N-1)$$

the second sum becomes

$$\sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{d_i d_j - NT_{i,j}}{N^2(N-1)} + \frac{4(N-1)\ell_G - N(N-1) - 4N\ell_G}{N^2(N-1)}.$$

Thus, we obtain

$$\mathbb{V}(H_{G^c}) = \mathbb{V}(H_G) + 1 - \frac{4\ell_G}{N(N-1)}$$

and, considering Theorem 3.3, the formula in (b) follows. \square

Some major remarks on the Theorem above are now in order.

REMARK 4.5. *As a first trivial example, we consider the limit situation of an empty graph. Let G be the empty graph. Its mp-chart has $\mathbb{E}(H_G) = \mathbb{V}(H_G) = 0$. In this case G^c is the complete graph with $N(N-1)/2$ edges and, by Theorem 4.4 part (b), one has $\mathbb{V}(H_{G^c}) = 1$. This can be verified also by direct computation. As a matter of fact, the adjacency matrix of G^c consists of non-zero entries out of the diagonal. Thus $d_i^c = N-1$ for all i and $T_{ij}^c = N-2$ for all i, j , with $i \neq j$. Hence*

$$\mathbb{V}(H_G) = \sum_{i=1}^N \frac{d_i(N-d_i)}{N^2} + \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{d_i d_j - N T_{i,j}}{N^2(N-1)} = \frac{N-1}{N} + \frac{(N-1)^2 - N(N-2)}{N^2(N-1)} = 1.$$

These computations show that the variance of the mp-chart of a complete graph is invariant on the number of vertices.

REMARK 4.6. *Few straightforward algebraic calculations show that the difference $\mathbb{V}(H_{G^c}) - \mathbb{V}(H_G)$ lies between -1 and 1 . Therefore, under the hypotheses of Theorem 3.8, when N goes to infinity we have that $\mathbb{V}(H_{G^c}) \cong \mathbb{V}(H_G)$. Intuitively, the difference between $\mathbb{V}(H_{G^c})$ and $\mathbb{V}(H_G)$ depends on the diagonal entries which are forced to be zero in the adjacency matrix A_{G^c} . The effect of these entries vanishes when the size of the graph goes to infinity.*

REMARK 4.7. *Another interesting property follows from Theorem 4.4. First, notice that $\mathbb{E}(H_{G^c}) = \mathbb{E}(H_G)$ implies that G and G^c have the same number of edges, namely $N(N-1)/4$. (This is not possible for all values of N). In such a case, H_{G^c} and H_G are forced to have the same variance, no matter how is complicated the graph G .*

The computation of the whole mp-chart of the complement graph G^c from the mp-graph of G is less easy. Given a graph G , we build a $(N+1) \times (N+1)$ -matrix M_G , indexed, both on rows and columns, by $\{0, \dots, N\}$, and defined as follows. The entry $(M_G)_{i,j}$ is the number of permutations π such that $\sum_{s=1}^N (A_G)_{s,\pi(s)} = j$ and π has i diagonal elements (that is, $\pi(s) = s$ for i elements).

EXAMPLE 4.8. *Consider the graph G with matrix*

$$A_G = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

The corresponding matrix M_G is

$$M_G = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 0 & 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix M_G allows the computation of both the mp-charts H_G and H_{G^c} . Roughly speaking, to compute the mp-chart of G is enough to sum the columns of M_G , while to compute the mp-chart of G^c we need to sum the entries of suitable diagonals of M_G . More precisely, the following relations hold true.

PROPOSITION 4.9. *For a graph G , we have for all $j = 1, \dots, N$:*

(a) *The components of the mp-chart H_G are:*

$$h_G(j) = \sum_{i=0}^N (M_G)_{i,j};$$

(b) *The components of the mp-chart H_{G^c} are:*

$$h_{G^c}(j) = \sum_{i=0}^{N-j} (M_G)_{i,N-j-i}.$$

Proof. The first relation follows by the definition of mp-chart, as the sum of entries in the j -th column of M_G is the number of permutations with j elements equal to 1, that is $h_G(j)$.

To prove the second relation, it is enough to prove that for all i and j we have: $(M_{G^c})_{i,j} = (M_G)_{i,N-j-i}$. Suppose that π is such that $\sum_{s=1}^N (A_G)_{s,\pi(s)} = j$ and π has i diagonal elements. When we consider π on $A(G^c)$, we have $A(G^c)_{s,\pi(s)} = 1$ for the s such that $(A_G)_{s,\pi(s)} = 0$, except for the diagonal entries where we still have 0. Hence there are $N - j - i$ entries in $A(G^c)$ such that $A(G^c)_{s,\pi(s)} = 1$. This completes the proof. \square

EXAMPLE 4.10. *Consider the complement graph G^c of the graph G in Example 4.8:*

$$A_{G^c} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Then

$$M_{G^c} = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 0 & 4 & 4 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We notice that in this simple case $M_G = M_{G^c}$. This is due to the fact that G and G^c , up to the labels of the vertices, are equivalent.

REMARK 4.11. *In principle, the max-plus permanent of the complement graph G^c can be computed from the matrix M_G . In fact, Proposition 4.9 shows that the mp-chart of the complement graph can be computed from the matrix M_G by adding along suitable diagonals and the max-plus permanent is just the position of the last non-zero element of the mp-chart. However, as the matrix M_G is not easy to compute for large graphs, this approach does not help in actual computations.*

5. Future directions. The max-plus permanent and the mp-chart studied in this paper lead to several new questions. In fact, we have analyzed here only the case of undirected graph. Therefore, among the future directions of our research, there will be the extension of the definition of max-plus permanent and mp-chart for undirected and weighted graphs. Moreover, special classes of graphs may be studied, such as bipartite graphs, or fixed-degree graphs. In particular, fixed-degree graphs correspond to adjacency matrices with fixed margins and in that context algebraic and combinatorial methods have demonstrated already their potential.

Another possible research direction is strictly in graph theory. As a matter of fact it could be interesting to compare the mp-chart of a graph with other well-known descriptors of its complexity. In a recent work in progress the mp-chart is compared to matching polynomials. By several examples we know that there exist different graphs with the same mp-chart. We notice that, in all these cases, also the matching polynomials coincide. So a principal question would be: If two non-isomorphic graphs have the same mp-chart, then their matching polynomials are equal?

Finally, applications to large graphs, possibly through simulation techniques, will be investigated in order to use the tools presented in this paper to real data examples.

REFERENCES

- [1] RÉKA ALBERT AND ALBERT-LÁSZLÓ BARABÁSI, *Statistical mechanics of complex networks*, Review of Modern Physics, 74 (2002), pp. 47–97.
- [2] ALEXANDER BARVINOK, *Enumerating contingency tables via random permanents*, Combin. Probab. Comput., 17 (2008), pp. 1–19.
- [3] ———, *On the number of matrices and a random matrix with prescribed row and column sums and 0–1 entries*, Adv. Math., 224 (2010), pp. 316–339.
- [4] ALEXANDER BARVINOK AND JOHN A. HARTIGAN, *The number of graphs and a random graph with a given degree sequence*. arXiv:1003.0356, 2010.
- [5] NORMAN BIGGS, *Algebraic Graph Theory*, Cambridge University Press, New York, 2 ed., 1993.
- [6] PATRICK BILLINGSLEY, *Probability and Measure*, John Wiley and Sons, New York, 3 ed., 1995.
- [7] BÉLA BOLLOBÁS, *Modern Graph Theory*, Springer-Verlag, New York, 1998.
- [8] ———, *Random Graphs*, Cambridge University Press, Cambridge, 2 ed., 2001.
- [9] ANIRBAN CHAKRABORTI, IOANE MUNI TOKE, MARCO PATRIARCA, AND FREDERIC ABERGEL, *Econophysics: Empirical facts and agent-based models*. arXiv:0909.1974, 2010.
- [10] REINHARD DIESTEL, *Graph Theory*, Springer, Heidelberg, 3 ed., 2005.
- [11] RICK DURRETT, *Random Graph Dynamics*, Cambridge University Press, New York, 2007.
- [12] UBALDO GARIBALDI AND ENRICO SCALAS, *Finitary Probabilistic Methods in Econophysics*, Cambridge University Press, Cambridge, UK, 2010.
- [13] STÉPHANE GAUBERT AND FRÉDÉRIC MEUNIER, *Carathéodory, Helly and the others in the max-plus world*, Discrete Comput. Geom., 43 (2010), pp. 648–662.
- [14] WASSILY Hoeffding, *A combinatorial central limit theorem*, Ann. Math. Statist., 22 (1951), pp. 558–566.
- [15] TAEHYUN HWANG, HUGUES SICOTTE, ZE TIAN, BAOLIN WU, JEAN-PIERRE KOCHER, DENNIS A. WIGLE, VIPIN KUMAR, AND RUI KUANG, *Robust efficient identification of biomarkers by classifying features on graphs*, Bioinformatics, 24 (2008), pp. 2023–2029.
- [16] ANDRÉ A. KELLER, *Graph theory and economic models: from small to large size applications*, Electronic Notes in Discrete Mathematics, 28 (2007), pp. 469–476.
- [17] OLIVER MASON AND MARK VERWOERD, *Graph theory and networks in biology*, IET Syst. Biol., 1 (2007), pp. 89–119.
- [18] MICHAEL B. MONAGAN, KEITH O. GEDDES, K. MICHAEL HEAL, GEORGE LABAHN, STEFAN M. VORKOETTER, JAMES MCCARRON, AND PAUL DEMARCO, *Maple 10 Programming Guide*, Maplesoft, Waterloo ON, Canada, 2005.
- [19] R DEVELOPMENT CORE TEAM, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2010. ISBN 3-900051-07-0.
- [20] ROBERT J. SERFLING, *Contribution to central limit theory for dependent variables*, Ann. Math. Statist., 39 (1968), pp. 1158–1175.
- [21] JONATHAN SILVER, ERIC SLUD, AND KEIJI TAKAMOTO, *Statistical equilibrium wealth distribu-*

tions in an exchange economy with stochastic preferences, Journal of Economic Theory, 106 (2002), pp. 417–435.