THE ENUMERATION OF VERTEX INDUCED SUBGRAPHS WITH RESPECT TO THE NUMBER OF COMPONENTS

P. TITTMANN, I. AVERBOUCH*, AND J.A. MAKOWSKY**

Abstract. Inspired by the study of community structure in connection networks, we introduce the graph polynomial $Q(G;x,y)$, the bivariate generating function which counts the number of connected components in induced subgraphs.

We give a recursive definition of $Q(G;x,y)$ using vertex deletion, vertex contraction and deletion of a vertex together with its neighborhood and prove a universality property. We relate $Q(G;x,y)$ to other known graph invariants and graph polynomials, among them partition functions, the Tutte polynomial, the independence and matching polynomials, and the universal edge elimination polynomial introduced by I. Averbouch, B. Godlin and J.A. Makowsky (2008).

We show that $Q(G;x,y)$ is vertex reconstructible in the sense of Kelly and Ulam, discuss its use in computing residual connectedness reliability. Finally we show that the computation of $Q(G;x,y)$ is $\sharp\mathbf{P}$-hard, but Fixed Parameter Tractable for graphs of bounded tree-width and clique-width.

Contents

Date: January 20, 2009.

* Partially supported by a grant of the Graduate School of the Technion–Israel Institute of Technology.

** Partially supported by a grant of the Fund for Promotion of Research of the Technion–Israel Institute of Technology and grant ISF 1392/07 of the Israel Science Foundation (2007-2010).

1. Introduction

1.1. **Motivation: Community Structure in Networks.** In the last decade stochastic social networks have been analyzed mathematically from various points of view. Understanding such networks sheds light on many questions arising in biology, epidemology, sociology and large computer networks. Researchers have concentrated particularly on a few properties that seem to be common to many networks: the small-world property, power-law degree distributions, and network transitivity, For a broad view on the structure and dynamics of networks, see [36]. M. Girvan and M.E.J. Newman, [24], highlight another property that is found in many networks, the property of *community structure*, in which network nodes are joined together in tightly knit groups, between which there are only looser connections.

Motivated by [35], and the first author's involvement in a project studying social networks, we were led to study the graph parameter $q_{ij}(G)$, the number of vertex subsets $X \subseteq V$ with i vertices such that $G[X]$ has exactly j components. $q_{ij}(G)$, counts the number of degenerated communities which consist of i members, and which split into j isolated subcommunities.

The ordinary bivariate generating function associated with $q_{ij}(G)$ is the two-variable graph polynomial

$$Q(G;x,y) = \sum_{i=0}^{n} \sum_{j=0}^{n} q_{ij}(G)\, x^i y^j .$$

We call $Q(G;x,y)$ the *subgraph component polynomial* of G. The coefficient of y^k in $Q(G;x,y)$ is the ordinary generating function for the number of vertex sets that induce a subgraph of G with exactly k components.

1.2. $Q(G;x,y)$ **as a Graph Polynomial.** There is an abundance of graph polynomials studied in the literature, and slowly there is a framework emerging, [30, 31, 25], which allows to compare graph polynomials with respect to their ability to distinguish graphs, to encode other graph polynomials or numeric graph invariants, and their computational complexity. In this paper we study the *s*ubgraph component polynomial $Q(G;x,y)$ as a graph polynomial in its own right and explore its properties within this emerging framework.

Like the bivariate Tutte polynomial, see [10, Chapter 10], the polynomial $Q(G;x,y)$ has several remarkable properties. However, its distinguishing power is quite different from the Tutte polynomial and other well studied polynomials.

Our main findings are:

- $Q(G;x,y)$ distinguishes graphs which cannot be distinguished by the matching polynomial, the Tutte polynomial, the characteristic polynomial, or the bivariate chromatic polynomial introduced in [18] (Section 3).
- Nevertheless, we construct an infinite family of pairs of graphs which cannot be pairwise distinguished by $Q(G;x,y)$ (Proposition 20).
- The Tutte polynomial, satisfies a linear recurrence relation with respect to edge deletion and edge contraction, and is universal in this respect. $Q(G;x,y)$ also satisfies a linear recurrence relation, but with respect to three kinds of vertex elimination, and is universal in this respect. (Theorems 13 and 21).

- A graph polynomial in three indeterminates, $\xi(G; x, y, z)$, which satisfies a linear recurrence relation with respect to three kinds of edge elimation, and which is universal in this respect, was introduced in [4, 5]. It subsumes both the Tutte polynomial and the matching polynomial. For a line graph $L(G)$ of a graph G, we have $Q(L(G); x, y)$ is a substitution instance of $\xi(G; x, y, z)$ (Theorem 22).
- For fixed positive integer n the univariate polynomial $Q(G; x, n)$ can be interpreted as counting weighted homomorphisms, [20], and is related to the Widom-Rowlinson model for n particles (Theorem 10).
- $Q(G; x, y)$ is reconstructible from its vertex deletion deck in the sense of [12, 11] (Theorem 27).
- $Q(G; x, y)$ can be used (Section 8), to compute the probability $P_k(G)$ that a vertex induced subgraph of G has exactly k components from the subgraph polynomial. For $k = 1$ this is known as the *residual connectedness reliability* (Section 8).
- Also like for the Tutte polynomial, cf. [27], $Q(G; x_0, y_0)$ has the *Difficult Point Property*, i.e. it is $\sharp\mathbf{P}$-hard to compute for all fixed values of $(x_0, y_0) \in \mathbb{R}^2 - E$ where E is a semi-algebraic set of lower dimension (Theorem 29). In [31] it is conjectured that the Difficult Point Property holds for a wide class of graph polynomials, the graph polynomials definable in Monadic Second Order Logic. The conjecture has been verified only for special cases, [6, 7, 8].
- $Q(G; x_0, y_0)$ is fixed parameter tractable in the sense of [19] when restricted to graphs classes of bounded tree-width (Proposition 31) or even to classes of bounded clique-width (Proposition 32). For the Tutte polynomial, this is known only for graph classes of bounded tree-width, [37, 2, 33].

Outline of the paper. The paper is organized as follows: In Section 2 we introduce the polynomial $Q(G; x, y)$ and its univariate versions. In Seection 3 we discuss the distinguishing power of $Q(G; x, y)$ and compare this to other graph polynomials. In Section 4 we show how certain graph parameters are definable using $Q(G; x, y)$ and relate it to partition functions and counting weighted homomorphisms. In Section 5 we give a recursive defintion of $Q(G; x, y)$ using deletion, contraction and extraction of vertices and show that $Q(G; x, y)$ is universal. We also compare it to the universal edge elimination polynomial $\xi(G; x, y, z)$ defined in [4, 5] and give a subset expansion formula for $Q(G; x, y)$. In Section 6 we prove decomposition formulas for $Q(G; x, y)$ for clique separators. In Section 7 we show the reconstructibility of $Q(G; x, y)$. In Section 8 we discuss its use to compute the residual connectedness reliability. In Section 9 we discuss the complexity of computing $Q(G; x, y)$. In Section 10 we draw conclusions and state open problems.

2. The Subgraph Component Polynomial $Q(G; x, y)$

2.1. **The Bivariate Polynomial.**

Let $G = (V, E)$ be a finite undirected graph with n vertices and let $k \leq n$ be a positive integer. Assume the vertices of G fail stochastic independently with a given probability $q = 1 - p$. What is the probability that a subgraph of G with exactly k components survives? The solution of this problem leads to the enumeration of vertex induced subgraphs of G with k components. For a given vertex subset $X \subseteq V$, let $G[X]$ be the *vertex induced subgraph* of G with vertex set X and all edges of G that have both end vertices in

X. We denote by $k(G)$ the number of components of G. Let $q_{ij}(G)$ be the number of vertex subsets $X \subseteq V$ with i vertices such that $G[X]$ has exactly j components:

$$q_{ij}(G) = |\{X \subseteq V : |X| = i \wedge k(G[X]) = j\}|$$

The ordinary generating function for these numbers is the two-variable polynomial

$$Q(G; x, y) = \sum_{i=0}^{n} \sum_{j=0}^{n} q_{ij}(G) x^i y^j.$$

We call $Q(G; x, y)$ the *subgraph component polynomial* of G. Since loops or parallel edges do not contribute to connectedness properties of a graph, we assume in this paper that all graphs are simple.

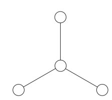

FIGURE 1. The star $Star_3 = K_{1,3}$

The star $K_{1,3}$, presented in Figure 1, has the subgraph polynomial

$$Q(K_{1,3}; x, y) = 1 + 4xy + 3x^2y + 3x^3y + x^4y + 3x^2y^2 + x^3y^3.$$

The term $3x^2y^2$ tell us that there are 3 possibilities to select two vertices of G that are non-adjacent.

The empty set induces the null graph $N = (\emptyset, \emptyset)$ that we consider as being connected, which gives $q_{00}(G) = Q(G; 0, 0) = 1$ for any graph. Substitution of 1 for y results in an univariate polynomial that is the ordinary generating function for all subsets of V, i.e. $Q(G; x, 1) = (1 + x)^n$.

2.2. **Univariate Polynomials.** The coefficient of y^k in $Q(G; x, y)$ is the ordinary generating function for the number of vertex sets that induce a subgraph of G with exactly k components:

$$Q_k(G; x) = \left[y^k\right] Q(G; x, y)$$

We call the polynomial Q_k for $k \in \mathbb{N}$ again *subgraph component polynomial.* The subscript as well as the number of variables should avoid confusion with the formerly defined subgraph polynomial. The subgraph polynomial $Q_1(G; x)$ is of special interest. We rename this polynomial to

$$S(G; x) = Q_1(G; x) = \sum_{i=0}^{n} s_i(G) x^i.$$

It counts the connected vertex induced subgraphs of G. A *separating vertex set* of a connected graph $G = (V, E)$ is a subset $X \subseteq V$ such that $G - X$ is a disconnected graph.

Theorem 1. *Let $G = (V, E)$ be a connected graph with n vertices. Let $c_k(G)$ be the number of separating vertex sets of cardinality k for $k = 0, 1, ..., n$. Then the coefficients of the subgraph polynomial $S(G; x)$ are given by*

$$s_k(G) = \binom{n}{k} - c_{n-k}(G).$$

Proof. If X is a separating vertex set then $V \setminus X$ induces a disconnected graph. Conversely, if $X \subseteq V$ is not a separating vertex set of G then $G[V \setminus X]$ is connected. □

We conclude that $2^n - S(G;1)$ is the number of all separating vertex sets of G.

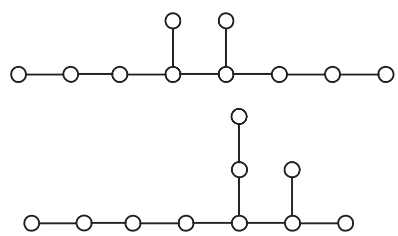

FIGURE 2. Non-isomorphic trees with the same subgraph polynomial

A graph invariant f is *trivial* on a class of graphs K of for any two graphs G_1 and G_2 with the same number of vertices we have $f(G_1) = f(G_2)$.

Proposition 2. *All non-isomorphic trees with up to nine vertices have different subgraph component polynomials. In other words we have: The graph polynomials* $Q_k(G;x)$ *and* $Q(G;x,y)$ *are not trivial on trees.*

However, we have:

Proposition 3. *There exist a unique pair of non-isomorphic trees with* 10 *vertices sharing the same subgraph component polynomial.*

Proof. Figure 2 shows these trees. This statement is true for $S(G;x)$ as well as for the (general) subgraph polynomial $Q(G;x,y)$. □

In Section 5 we shall see how to use this to generate infinite families of pairs of graphs which are not distinguished by $Q(G;x,y)$.

3. Distinctive Power

We denote by $m(G;x) = \sum_i m_i(G)x^i$ be the matching polynomial with $m_i(G)$ the number of i-matchings of G, by $p(G;x)$ be the characteristic polynomial, by $T(G;x,y)$ the Tutte polynomial, and by $P(G;x,y)$ the bivariate chromatic polynomial introduced in [18].

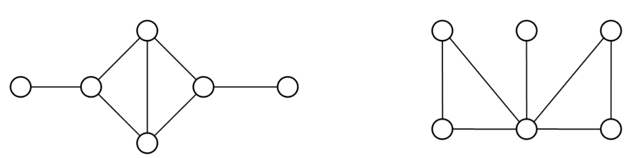

FIGURE 3. The graphs G_1, G_2

FIGURE 4. The graphs G_3, G_4

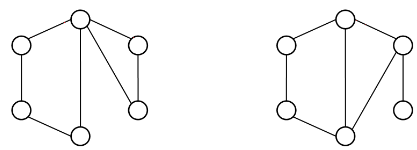

FIGURE 5. The graphs G_5, G_6

Proposition 4. *For the graphs $G_i; i = 1, \dots 6$ from Figures 3, 4, 5, and for P_4 and $K_{1,3}$ we have*

(1) $p(G_1; x) = p(G_2; x)$ *but* $Q(G_1; x, y) \neq Q(G_2; x, y)$.
(2) $m(G_3; x) = m(G_4; x)$ *but* $Q(G_3; x, y) \neq Q(G_4; x, y)$.
(3) $P(G_5; x, y) = P(G_6; x, y)$ *but* $Q(G_5; x, y) \neq Q(G_6; x, y)$.
(4) $T(P_4; x, y) = T(K_{1,3}; x, y)$ *but* $Q(P_4; x, y) \neq Q(K_{1,3}; x, y)$.

Proof. (1) and (2) are easy to verify.
For (3) $P(G_5; x, y) = P(G_6; x, y)$ is from [18]. For $Q(G_5; x, y) \neq Q(G_6; x, y)$ we compare $[x^4y^3]Q(G_5, x, y)$ with $[x^4y^3]Q(G_6, x, y)$.
For (4) we use that the Tutte polynomial does not distinguish trees of the same size, but that $Q(G; x, y)$ distinguishes all trees of size up to nine vertices, 2. □

Remark 5. *$Q(G; x, y)$ does not distinguish between graphs which differ only by the multiplicity of their edges, whereas for the Tutte polynomial this is not the case. Let $K_n^{(m)}$ denote the complete graph with m edges between any two distinct vertices. Then we have $T(K_n^1; x, y) \neq T(K_n^2; x, y)$ but $Q(K_n^1; x, y) = Q(K_n^2; x, y)$.*

Problem 6. *Are there simple graphs distinguished by $p(G; x)$, $m(G; x)$, $P(G; x, y)$ or $T(G; x, y)$ which are not distinguished by $Q(G; x, y)$?*

We say that a simple graph G is *determined by a graph polynomial f* if for every simple graph G' such that $f(G) = f(G')$ we have that G is isomorphic to G'. The class of simple graphs K is *determined by a graph polynomial f* if every graph $G \in K$ is determined by f. This notion has been studied in [38, 16], for the chromatic polynomial, the Tutte polynomial and the matching polynomial. It is shown, e.g., that several well-known families of graphs are determined by their Tutte polynomial, among them the class of wheels, squares of cycles, complete multipartite graphs, ladders, Möbius ladders, and hypercubes.

It follows from Proposition 8 that the class of empty graphs E_n is determined by $Q(G; x, y, z)$, and so is the class of complete graphs K_n. Note that, since $T(E_n; x, y) = 1$ for all $n \in \mathbb{N}$, the class of empty graphs is not determined by the the Tutte polynomial. It follows from Proposition 3 that the class of trees is not determined by $Q(G; x, y, z)$.

Proposition 7. *The class of graphs of the form $Star_n = K_{1,n}$ is determined by $Q(G; x, y, z)$.*

Proof. It is easy to verify, that if $[x^{n+1}y]Q(G; x, y) = 1$, $[x^n y^n]Q(G; x, y) = 1$ and $[x^{n+2}]Q(G; x, y) = 0$, then G is isomorphic to $Star_n$. □

4. Combinatorial Interpretations

4.1. Evaluations and Coefficients of $Q(G; x, y)$.

For a polynomial $f(x, y)$, let $\left[x^i y^j\right] f(x, y)$ be the coefficient of $x^i y^j$ in $f(x, y)$ and let $\deg_x f$ be the degree with respect to the variable x.

Proposition 8. *The following graph properties can be easily obtained from the subgraph polynomial:*

(1) *The number of vertices:*

$$n(G) = |V(G)| = \deg_x Q(G; x, y) = \log_2 Q(G; 1, 1)$$

(2) *The number of edges:*

$$e(G) = \left[x^2 y\right] Q(G; x, y)$$

(3) *The number of components:*

$$k(G) = \deg\left(\left[x^{n(G)}\right] Q(G; x, y)\right)$$

Theorem 9. *The degree of the subgraph component polynomial $Q(G; x, y)$ with respect to y is the cardinality of a maximum independent set of G (the independence number):*

$$\deg_y Q(G; x, y) = \alpha(G)$$

Proof. Let $X \subseteq V$ be a maximum independent set of G. In this case, we have $k(G[X]) = |X|$ and hence $\deg_y Q(G; x, y) \geq \alpha(G)$. Assume that there exists a set $Y \subseteq V$ with $k(G[Y]) > |X|$. Then we obtain an independent set X' by selecting one vertex of each component of $G[Y]$ such that $|X'| > |X|$ – a contradiction. □

Let $a_i(G)$ be the number of independent vertex sets of size i of G. The *independence polynomial* of G is defined by

$$I(G; x) = \sum_{i=0}^{n} a_i(G) x^i.$$

As a consequence of Theorem 9 we can derive the independence polynomial of G from the subgraph component polynomial. Let $\left[y^j\right] Q(G; x, y)$ denote the coefficient of y^j in $Q(G; x, y)$. This coefficient is a polynomial in x where the coefficient of x^i counts the vertex subsets of cardinality i of G that induce a subgraph with j components. A vertex set $X \subseteq V$ is independent if and only if $k(G[X]) = |X|$. Hence $\left[x^j\right]\left[y^j\right] Q(G; x, y) = a_j(G)$ is the number of independent vertex sets of size j of G.

4.2. **Partition Functions.** In this subsection we show that the subgraph polynomial $Q(G; x, y_0)$ for any $x \in \mathbb{R}$ and fixed $y_0 \in \mathbb{N}$ can be viewed as a partition function, using counting of weighted graph homomorphisms. Partition functions were first studied in the context of statistical physics and have recently attracted much attention, [34]. A systematic study of the question which graph invariants can be presented as partition functions has been initiated in [23].

A weighted graph (H, α, β) consists of a graph $H = (V(H), E(H))$ with α assigning weights to vertices and β to edges. The partition function $Z_{H,\alpha,\beta}(G)$ associated with (H, α, β) is defined by

$$Z_H(G) = \sum_{\substack{h\,:\,V \mapsto V_H \\ homomorphism}} \prod_{v \in V} \alpha(h(v)) \prod_{(u,v)\in E} \beta(h(u), h(v))$$

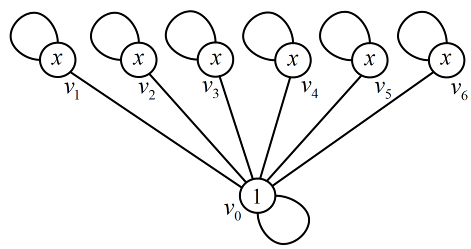

FIGURE 6. Auxiliary graph $Star_y$ for $y = 6$

Let $(Star_y, \alpha, \beta)$ be a weighted star with $y+1$ vertices and with all loops. The central vertex is v_0. An example of $Star_y$ for $y = 6$ is shown on Figure 6. The weight functions are defined as follows:

$$\alpha(v) = \begin{cases} 1 & if\ v = v_0 \\ x & otherwise \end{cases}$$

$$\beta(u, v) = 1$$

Theorem 10. *Let $Z_{(H,\alpha,\beta)}(G)$ be the partition function associated with $H = Star_y$ and α, β as above. Then, for all nonnegative integers y and all $x \in \mathbb{R}$, we have*

$$Q(G; x, y) = Z_{(Star_y,\alpha,\beta)}$$

Proof. Let us start with the definition of $Z_H(G)$. Under every mapping $h : V \mapsto V_H$, let $A \subseteq V$ be the subset of vertices that are not mapped to v_0. Let us count the homomorphisms that map the subset A into $v_1, \ldots, v_y$: there are exactly $y^{k(G[A])}$ such homomorphisms, because every connected component of $G[A]$ must be mapped into a single vertex. Finally, we get

$$\begin{aligned} Z_H(G) &= \sum_{\substack{h\,:\,V \mapsto V_H \\ homomorphism}} \prod_{v\in V} \alpha(h(v)) = \\ = \sum_{\substack{h\,:\,V \mapsto V_H \\ homomorphism}} x^{|A|} &= \sum_{A\subseteq V} \sum_{\substack{h\,:\,V \mapsto V_H \\ homomorphism \\ v\in A \leftrightarrow h(v)\neq v_0}} x^{|A|} = \\ = \sum_{A\subseteq V} y^{k(G[A])} x^{|A|} &= Q(G;x,y) \end{aligned}$$

which by (8) completes the proof □

It is open whether there are other points in which $Q(G;x,y)$ is definable as a partition function.

Remark 11. *The auxiliary graph $H = Star_n$ is called in physical literature* The Widom-Rowlinson model for n particles*. The homomorphisms to $Star_n$ are called* Widom-Rowlinson configurations*,* [20]*.*

5. Recursive Definition and Subset Expansion

5.1. **Recurrence Relation for Vertex Elimination.** We turn now our attention to the investigation of properties of the subgraph polynomial that support its computation. The first statement concerns the multiplicativity with respect to components of the graph.

Theorem 12 (Multiplicativity)**.** (1) *Let $G = G_1 \sqcup G_1$ be the disjoint union of the graphs G_1 and G_2. Then*

$$Q(G;x,y) = Q(G_1;x,y)\cdot Q(G_2;x,y).$$

(2) *In particular, if $G = (V,E)$ consists of c components $G_1,...,G_c$ then the subgraph polynomial satisfies*

$$Q\left(G;x,y\right) = \prod_{j=1}^{c} Q\left(G_j;x,y\right).$$

Proof. In case $c = 2$, each subset $X \subseteq V$ of cardinality k is the disjoint union of two subsets $X_1 \subseteq V\left(G_1\right)$ and $X_2 \subseteq V\left(G_2\right)$ with $|X_1| = j$ and $|X_2| = i-j$. The number of components of $G\left[X\right] = G\left[X_1 \cup X_2\right]$ is the sum of the number of components of $G\left[X_1\right]$ and $G\left[X_2\right]$. We obtain

$$q_{ik}\left(G\right) = \sum_{j=0}^{i}\sum_{l=0}^{k} q_{jl}\left(G_1\right) q_{i-j,k-l}\left(G_2\right). \tag{1}$$

Thus for a graph with two components, the subgraph polynomial satisfies

$$Q\left(G;x,y\right) = Q\left(G_1;x,y\right) Q\left(G_2;x,y\right).$$

We obtain the statement of the theorem by induction on the number of components. □

We distinguish three types of *vertex elimination*:

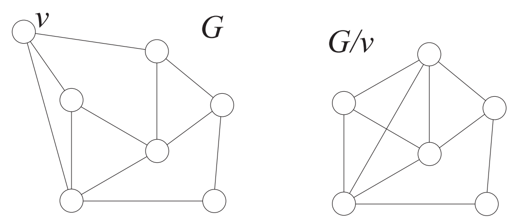

Figure 7. Vertex contraction

Vertex deletion: For a given vertex $v \in V(G)$, let $G-v$ the graph obtained from G by removal of v and all edges that are incident to v. We call this operation *vertex deletion*.

Vertex extraction: Similarly, let $G-X$ be the graph obtained from G by removal of all vertices of the set $X \subseteq V$. Let $N(v)$ be the set of vertices that are adjacent to v in G (the neighborhood of v). We denote by $N[v]$ the *closed neighborhood* of a vertex v in G, i.e. the set of all vertices adjacent to v including v itself. The operation $G-N[v]$ is called *vertex extraction.*

Vertex contraction: A further special graph operation is needed here – the *contraction* of a vertex. That is the graph G/v obtained from G by removal of v and insertion of edges between all pairs of non-adjacent neighbor vertices of v. Figure 7 shows an example graph and the graph obtained by vertex contraction.

Theorem 13. *Let $G=(V,E)$ be a graph and $v \in V$. Then the subgraph polynomial satisfies the decomposition formula*

$$Q(G;x,y) = Q(G-v;x,y) + x(y-1)\,Q(G-N[v];x,y) + xQ(G/v;x,y).$$

Proof. Let us first consider all vertex induced subgraphs of G that do not contain vertex v. These subgraphs are also vertex induced subgraphs of $G-v$. Consequently,

$$Q(G-v;x,y)$$

enumerates all induced subgraphs not including the vertex v.

In a second step we count all vertex induced subgraphs that contain vertex v but none of its neighbors in G. In this case, the vertex v forms a connected component consisting of v only. The rest of the induced subgraph is a subgraph of of $G-N[v]$. All these subgraphs are enumerated by $Q(G-N[v];x,y)$. However, the component built by v contributes one vertex and one component to the polynomial. Thus we obtain the generating function

$$xyQ(G-N[v];x,y).$$

In our enumeration so far we missed exactly those vertex induced subgraphs that contain v and at least one of its neighbors together in one component. We include v in the corresponding candidate set, remove it from G, and multiply the generating function by x (not by xy because we do not increase the number of components). In order to trace the components, we have to simulate the paths using v in G. These paths are no longer present in $G-v$. This task is best performed by using G/v instead of $G-v$. Thus we obtain the contribution $xQ(G/v;x,y)$ to the generating function. Unfortunately, this polynomial enumerates induced subgraphs that do not contain any vertices from $N(v)$, too. We can fix this problem by subtraction of $xQ(G-N[v];x,y)$, which gives

$$xQ(G/v;x,y) - xQ(G-N[v];x,y)$$

as final contribution to the generating function. □

Corollary 14. *Let $v \in V$ be a vertex of degree 1 in $G=(V,E)$ and let w be its only neighbor in G. Then*

$$Q(G;x,y) = (1+x)\,Q(G-v;x,y) + x(y-1)\,Q(G-\{v,w\};x,y).$$

Proof. Notice that in this case $G/v = G-v$ and $G-N[v] = G-\{v,w\}$. Then the statement follows immediately from Theorem 13. □

5.2. **Some Easy Computations.** The subgraph polynomial can be easily computed for certain special graphs.

Proposition 15. *For the complete graphs K_n and the empty graphs E_n (the complement of K_n) we have:*

(1) $Q(K_n;x,y) = y(1+x)^n - y + 1$.
(2) $Q(E_n;x,y) = (1+xy)^n$.

Proof. (1) In a complete graph K_n each vertex subset except the empty set induces a connected subgraph.

(2) In the empty graph E_n each subset $X \subseteq V$ induces a subgraph with $|X|$ components. □

From Theorem 13 we obtain a recurrence relation for the subgraph polynomial of the paths P_n:

Proposition 16.

$$Q(P_n;x,y) = (1+x)\,Q(P_{n-1};x,y) + x(y-1)\,Q(P_{n-2};x,y)$$

Together with the initial values,

$$\begin{aligned} Q(P_0;x,y) &= 1, \\ Q(P_1;x,y) &= 1+xy, \end{aligned}$$

equation (16) determines the subgraph polynomial of P_n uniquely. The explicit solution is

$$Q(P_n;x,y) = \frac{1-x+a}{2a}\left(\frac{2x(1-y)}{1+x-a}\right)^{n+1} - \frac{1-x-a}{2a}\left(\frac{2x(1-y)}{1+x+a}\right)^{n+1}$$

with $a = \sqrt{1-2x+x^2+4xy}$.

The subgraph polynomial of the cycle C_n satisfies another recurrence relation:

Proposition 17.

$$Q(C_n;x,y) = Q(P_{n-1};x,y) + x(y-1)\,Q(P_{n-3};,x,y) + xQ(C_{n-1};x,y)$$

The *join* $G \vee H$ of two graphs $G=(V,E)$ and $H(W,F)$ with $V \cap W = \emptyset$ is the graph obtained from $G \cup H$ by introducing edges form each vertex of G to each vertex of H. Consequently, the join of two empty graphs $\overline{K_s}$ and $\overline{K_t}$ is the complete bipartite graph $K_{s,t}$.

Theorem 18. *Let $G = (V, E)$ and $H(W, F)$ be two graphs with $V \cap W = \emptyset$, $|V| = s$, $|W| = t$. Then*

$$Q(G \vee H; x, y) = Q(G; x, y) + Q(H; x, y) + [(1+x)^s - 1]\left[(1+x)^t - 1\right] y - 1.$$

Proof. All vertex subsets of $V \cup W$ belong to exactly one of three classes:
(1) subsets of V,
(2) subsets of W,
(3) subsets that have at least one vertex of V and at least one vertex of W.

The first two classes are counted by $Q(G; x, y)$ and $Q(H; x, y)$, respectively. The empty set is counted twice, which is corrected by subtracting one. All vertex subsets of the third class induce connected subgraphs of $G \vee H$. The generating function for the number of subsets of this class is $[(1+x)^s - 1]\left[(1+x)^t - 1\right]$. □

From Theorem 18, we deduce the subgraph polynomial of the complete bipartite graph:

Corollary 19.

$$Q(K_{s,t}; x, y) = (1+xy)^s + (1+xy)^t + [(1+x)^s - 1]\left[(1+x)^t - 1\right] y - 1$$

Propositions 16 and 17 and Theorem 18 are explicit instances of general results which follow from the fact that $Q(G; x, y)$ is definable in Monadic Second Order Logic. We shall discuss this feature in Section 5.5.

We can use Theorem 18 and the multiplicativity of $Q(G; x, y)$ to prove the following:

Proposition 20. *There are infinite families of pairs of non-isomorphic graphs with a fixed number of connected components which are not distinguished by $Q(G; x, y)$.*

Proof. Let G be a graph. We define inductively

$$\begin{aligned} F_0(G) &= G \\ J_0(G) &= G \\ F_{n+1}(G) &= F_n(G) \sqcup G \\ J_{n+1}(G) &= J_n(G) \vee G \end{aligned}$$

Let Tr_1 and Tr_2 be the two trees from Figure 2. Then, using Theorem 18 and the multiplicativity of $Q(G; x, y)$ we have for all $n \in \mathbb{N}$

$$Q(F_n(Tr_1); x, y) = Q(F_n(Tr_2); x, y)$$

and

$$Q(J_n(Tr_1); x, y) = Q(J_n(Tr_2); x, y)$$

For G connected, the graphs $J_n(G)$ are connected. The graphs $F_n(G)$ have exactly n components. So for m components we combine $F_{m-1}(G) \sqcup J_n(G)$ which has m components. □

5.3. **The Universality Property of $Q(G; x, y)$.** The vertex decomposition formula represented in Theorem 13 can be considered as a vertex equivalent to the well-known edge decomposition (deletion-contraction relations). Edge decomposition formulae of the form $f(G) = \alpha(e) f(G - e) + \beta(e) f(G/e)$ apply to the Tutte polynomial and derived graph invariants, for instance the number of spanning trees or the reliability polynomial. Indeed, it was shown by J.G. Oxley and D.J.A. Welsh,

[41], that the Tutte polynomial is in a certain sense *universal*, meaning that all other graph invariants that satisfy edge decomposition formulae can be derived from the Tutte polynomial by substitution of variables. A textbook presentation is given in [10]. A general framework analyzing universality properties of graph polynomials is studied in [25].

It seems natural to ask for the most general vertex decomposition formula. Let us assume that we try to construct an ordinary generating function $f(G)$ that counts some type of vertex induced subgraphs with respect to the number of vertices. Which properties should such a function have? If the subgraphs in question are composed from subgraphs of the components then we can expect multiplicativity of f with respect to components of the graph. In order to assign the value $f(G)$ uniquely to a graph G by application of a decomposition formula as given in Theorem 13, certain initial values for the null graph and the empty graph have to be given. Therefore, we presuppose the following four properties of f:

(a) (Multiplicativity) If G_1 and G_2 are components of G then $f(G) = f(G_1) f(G_2)$.
(b) (Recurrence relation) Let $\alpha, \beta, \gamma \in \mathbb{R}$ and let v be a vertex of G, then

$$f(G) = \alpha f(G - v) + \beta f(G - N[v]) + \gamma f(G/v). \tag{2}$$

(c) (Initial condition) There exists $\delta \in \mathbb{R}$ such that $f(\emptyset) = \delta$ for the null graph $\emptyset = (\emptyset, \emptyset)$.
(d) (Initial condition) There exists $\varepsilon \in \mathbb{R}$ such that $f(E_1) = \varepsilon$ for a graph $E_1 = (\{v\}, \emptyset)$ consisting of one vertex.

Furthermore, in order to make f a well-defined graph polynomial, the result of computing f has to be the same, irrespective of the order in which we apply the enabled computation steps. In particular, it has to be *independent of the order of the vertices*, which we use to apply the decomposition formula (b). In general we may choose $\alpha, \beta, \gamma, \delta, \varepsilon$ from a field of characteristic zero or from a ring. A graph invariant is *proper* if there are two graphs G_1 and G_2 with the same number of vertices such that $f(G_1) \neq f(G_2)$.

Applying the conditions (b), (c), and (d) we obtain from $E_1 - v = E_1 - N[v] = E_1/v = \emptyset$ the equation

$$\varepsilon = (\alpha + \beta + \gamma)\,\delta.$$

Computing $f(E_2) = f(\overline{K_2})$ in two ways using (a) and (b), respectively, results in

$$\varepsilon^2 = (\alpha + \beta + \gamma)\,\varepsilon.$$

Consequently, the values of δ and ε are determined:

$$\begin{aligned} \delta &= 1 \\ \varepsilon &= \alpha + \beta + \gamma \end{aligned}$$

If the constants α, β, γ are properly defined then the value of $f(G)$ does not depend on the choice of the vertex v in equation (2). Consequently, the function $f(G)$ does not depend on the order of the vertex decomposition (2). The calculation of $f(P_3)$ for path of three vertices yields, in case we start from a vertex of degree 1,

$$f(P_3) = (\alpha + \gamma)^2 (\alpha + \beta + \gamma) + \beta(\alpha + \gamma) + \beta(\alpha + \beta + \gamma).$$

If we begin the vertex decomposition at the vertex of degree 2 then we obtain

$$f(P_3) = \alpha(\alpha + \beta + \gamma)^2 + \beta + \gamma(\alpha + \gamma)(\alpha + \beta + \gamma) + \beta\gamma$$

These two results coincide if $\beta = 0$, $\alpha = 1$, or $\alpha + \beta + \gamma = 1$. In any case, there remain only two variables that can be chosen independently. In case of $\beta = 0$, all graphs with the same number of vertices result in the same polynomial. Therefore, this case does not yield any interesting applications. If $\alpha+\beta+\gamma = 1$ then $f(G) = 1$ for all graphs. The only remaining choice, $\alpha = 1$, gives for $\beta = x(y-1)$ and $\gamma = x$ the subgraph component polynomial $Q(G; x, y)$.

From this we get that $Q(G; x, y)$ is universal among polynomials recursively defined using vertex deletion, vertex extraction and vertex contraction. More precisely, we have the following theorem.

Theorem 21 (Universality of $Q(G; x, y)$)**.** (1) *For a graph polynomial $f(G; \alpha, \beta, \gamma, \delta, \varepsilon)$ to be proper and well-defined by conditions (a)-(d) we have $\alpha = 1$, $\delta = 1$ and $\varepsilon = 1 + \beta + \gamma$.*

(2) *There is a unique proper graph polynomial $U(G; \beta, \gamma)$ which is well-defined by conditions (a)-(d) and we have*

$$Q(G; x, y) = U(G; x(y-1), x)$$

and

$$U(G; \beta, \gamma) = Q(G; \gamma, \frac{\beta}{\gamma} + 1)$$

5.4. **Vertex Eliminations vs Edge Elimination.** The subgraph component polynomial $Q(G; x, y)$ can be regarded as counting vertex set expansions. In the literature there is a variety of graph polynomials, including the Tutte polynomial, which can be defined by counting edge set expansions.

We have seen in Theorem 21 that $Q(G; x, y)$ is universal among the polynomials defined recursively via deletion, extraction and contraction of vertices. In [4, 5] the polynomial $\xi(G; x, y, z)$ was shown to be universal among the polynomials defined recursively via deletion, extraction and contraction of edges. In this section we will show the connection of $G(G; x, y)$ to the universal edge elimination polynomial $\xi(G; x, y, z)$.

The polynomial $\xi(G; x, y, z)$ generalizes both the Tutte and the matching polynomials, as well as the bivariate chromatic polynomial of [18]. We shall use the recursive decomposition of $\xi(G; x, y, z)$ from [5]:

$$\begin{aligned}
&\xi(G; x, y, z) = \xi(G - e; x, y, z) + y\xi(G/e; x, y, z) + z\xi(G \dagger e; x, y, z) \\
&\xi(G_1 \sqcup G_2; x, y, z) = \xi(G_1; x, y, z)\xi(G_2; x, y, z) \\
&\xi(E_1; x, y, z) = x \\
&\xi(\emptyset) = 1 \qquad (3)
\end{aligned}$$

where $G_1 \sqcup G_2$ denotes the disjoint union of graphs G_1 and G_2, and the three edge elimination operations are defined as follows:

- **Edge deletion::** We denote by $G - e$ the graph obtained from G by simply removing the edge e.
- **Edge extraction::** We denote by $G \dagger e$ the graph induced by $V \setminus \{u, v\}$ provided $e = \{u, v\}$. Note that this operation removes also all the edges adjacent to e.
- **Edge contraction::** We denote by G/e the graph obtained from G by unifying the endpoints of e.

We will rewrite the decomposition of $Q(G;x,y)$ using Theorem 13.

$$\begin{aligned} &Q(G;x,y) = Q(G-v;x,y) + xQ(G/v;x,y) + x(y-1)Q(G-N[v];x,y) \\ &Q(G_1 \sqcup G_2;x,y) = Q(G_1;x,y)Q(G_2;x,y) \\ &Q(E_1;x,y) = xy+1 \\ &Q(\emptyset) = 1 \end{aligned} \tag{4}$$

Theorem 22. *Let $G = (V,E)$ be a graph. Let $L(G) = (V_e, E_e)$ denote the line graph of G. Then the following equation holds:*

$$\xi(G;1,x,x(y-1)) = Q(L(G);x,y)$$

Proof. First, let us analyze the correspondence of the edge elimination operations in a graph to the vertex elimination operations in its line graph. Let $v_e \in V_e$ be the vertex in the line graph that corresponds to the edge $e \in E$ of the original graph. By the definition of the edge and vertex elimination operations:

$$L(G-e) = L(G) - v_e \tag{5}$$

$$L(G/e) = L(G)/v_e \tag{6}$$

$$L(G \dagger e) = L(G) - N[v_e] \tag{7}$$

First let us check the connected graphs with up to one edge:

If $G \in \{\emptyset, E_1\}$, $L(G) = \emptyset$,
The equivalence $\xi(G;1,x,x(y-1)) = 1 = Q(\emptyset)$ holds.

If G is a single point with a loop, or $G = P_2$, $L(G)$ is a singleton, The equivalence $\xi(G;1,x,x(y-1)) = 1 + x + x(y-1) = 1 + xy = Q(E_1)$ holds.

Next, we note that $L(G_1 \sqcup G_2) = L(G_1) \sqcup L(G_2)$. Therefore, if the theorem holds for graphs G_1 and G_2, then it holds also for $G_1 \sqcup G_2$. Finally, the theorem follows by induction on the number of edges, using the decomposition formulae (4) and (3) and the correspondence of edge and vertex elimination operations. □

Problem 23. *How does the distinguishing power of $\xi(G;x,y,z)$ compare to the distinguishing power of $Q(G;x,y)$?*

5.5. **Subset Expansion and Definability in Logic.** $Q(G;x,y)$ was defined as a generating function. Let us rewrite the definition of $Q(G;x,y)$ in a slightly different way. Instead of summation over the number of the used vertices i , and the number of induced connected components j, we shall summate over all the possible subsets of vertices:

$$Q(G;x,y) = \sum_{A \subseteq V} x^{|A|} y^{k(G[A])}. \tag{8}$$

This is a *subset expansion formula*, a term coined in [42]. The relationship between recursive definitions of graph polynomials and the existence of subset expansion formulas has been studied from a logical point of view in [25]. Subset expansion formulas can often be used to show that a graph polynomial is definable in Monadic Second Order Logic, as studied in [28, 31] . However, the exponent $k(G[A])$ in Equation (8) causes a problem. to remedy this, we use, like in [29], an auxiliary order $\prec$ over the vertices. We will denote by $F(A)$ the subset of the *smallest* vertices in every respective connected component.

$$Q(G;x,y)=\sum_{A\subseteq V}\left(\prod_{v\in A}x\right)\left(\prod_{u\in F(A)}y\right) \tag{9}$$

Note that the result does not depend on the used auxiliary order.

Without having to go in the details of graph polynomials definable in Monadic Second Order Logic[1], Equation (9) shows that $Q(G;x,y)$ is a graph polynomial definable in Monadic Second Order Logic for graphs $G=(V,E)$ with universe V and a binary edge relation. Therefore all the theorems from [28, 32] can be applied. In particular, the Feferman-Vaught-type theorems from [28] guarantee existence of reduction formulas like multiplicativity from Theorem 12, or the one in Theorem 18 for the join, not only for the disjoint union or the join operation, but for a wide class of **MSOL**-definable operations. Also, a general theorem from [32] guarantees the existence of recurrence formulas, as proven in Propositions 16 and 17, for a wide class of recursively defined families of graphs, as studied also in [39]. Among these we have the wheels W_n, the ladders L_n and the stars $Star_n$. It should not be difficult to compute the recurrence relations for these explicitly.

We shall exploit **MSOL**-definability also for our complexity analysis in Section 9.2.

6. Clique Separators

The simplest case of a clique separator in a graph G is an *articulation*, i.e. a vertex whose removal from G results in an increase of the number of components. Let v be an articulation of G and let H and K be subgraphs of G such that $G=H\cup K$ and $H\cap K=(\{v\},\emptyset)$. It is well known, cf. [10], that in this case the Tutte polynomial $T(G;x,y)$ satisfies

$$T(G;x,y)=T(H;x,y)\cdot T(H;x,y)$$

In the case of the subgraph component polynomial the situation is a bit more complicated:

Theorem 24. *Let v be an articulation of G and let H and K be subgraphs of G such that $G=H\cup K$ and $H\cap K=(\{v\},\emptyset)$. Then the subgraph polynomial $Q(G)=Q(G;x,y)$ satisfies*

$$\begin{aligned}Q(G) &= Q(H-v)\,Q(K-v)\\ &\quad+\frac{1}{xy}\left[Q(H)-Q(H-v)\right]\left[Q(K)-Q(K-v)\right].\end{aligned}$$

Proof. The first product of the polynomial is the generating function for the number all vertex induced subgraphs that do not contain the articulation v. The product is justified by Theorem 12. The second term counts all remaining subgraphs, i.e. those ones containing vertex v. Here the equation (1) from the proof of Theorem 12 has to be modified. The vertex v is counted twice because it belongs to both K and H. This double counting is corrected by multiplication with x^{-1}. By analogy, we introduce the factor y^{-1} in order to avoid that the component containing v is counted twice. □

[1]The interested reader can consult [21] for the use Monadic Second Order Logic in finite model theory, and [13] for its use in graph theory.

Theorem 24 can be generalized in order to cover clique separators with more than one vertex. Let $G=(V,E)$ be a connected graph and H, K subgraphs of G such that $H\cap K=K_r$ and $H\cup K=G$. In this case $K_r=(U,F)$ forms a *separating clique* of G. Here we assume that neither H nor K coincides with K_r. The subgraphs H and K are called *split components* of G with respect to K_r.

Theorem 25. *Let $K_r=(U,F)$ be a clique separator of G such that there are two split components H and K. Then the subgraph polynomial $Q(G)=Q(G;x,y)$ satisfies*

$$\begin{aligned} Q(G) &= Q(H-U)\,Q(K-U) \\ &\quad+\frac{1}{y}\sum_{\emptyset\neq A\subseteq U}\frac{1}{x^{|A|}}\sum_{B\supseteq U\setminus A}\sum_{C\supseteq U\setminus A}(-1)^{|B|+|C|}\,Q(H-B)\,Q(K-C)\,. \end{aligned}$$

Proof. First we count all subgraphs that are induced by vertex subsets included in $V\setminus U$. These subgraphs are also subgraphs of $G-U$. From Theorem 12 we obtain $Q(H-U)\,Q(K-U)$ as generating function for all subgraphs of G induced by subsets of $V\setminus U$.

For each subset $A\subseteq U$, let $f_{ij}(H,A)$ be the number of vertex subsets $X\subseteq V(H)$ of cardinality i with $A\subseteq X$ such that the induced subgraph $H[X]$ has exactly j components:

$$f_{ij}(H,A)=\left|\{X: A\subseteq X\subseteq V(H)\wedge |X|=i\wedge k(H[X])=j\}\right|$$

The polynomial

$$F(H,A)=\sum_{i=0}^{n}\sum_{j=0}^{n}f_{ij}(H,A)\,x^i y^j$$

is the ordinary generating function for the numbers $f_{ij}(H,A)$. We define the numbers $f_{ij}(K,A)$ and the corresponding generating function $F(K,A)$ for the second split component analogously. Let $X\subseteq V(G)$ be a vertex subset with $X\cap U=A$. Then the component of $G[X]$ that contains A is counted in $F(H,A)$ and in $F(K,A)$. There is indeed only one component counted twice, since A induces a clique of H and K, respectively. Thus we obtain

$$Q(G)=Q(H-U)\,Q(K-U)+\frac{1}{y}\sum_{\emptyset\neq A\subseteq U}\frac{1}{x^{|A|}}F(H,A)\,F(K,A)\,. \tag{10}$$

The factor $x^{-|A|}$ takes into account that all vertices of A contribute to $F(H,A)$ and to $F(K,A)$. For each subset $B\subseteq U$, the subgraph polynomial of $H-B$ can be represented as a sum of generating functions as follows:

$$Q(H-B)=\sum_{A\subseteq U\setminus B}F(H,A)$$

We define $\hat{Q}(H,U\setminus B)=Q(H-B)$ and obtain

$$\hat{Q}(H,U\setminus B)=\sum_{A\subseteq U\setminus B}F(H,A)$$

or

$$\hat{Q}(H,B)=\sum_{A\subseteq B}F(H,A)\,.$$

By Möbius inversion, we obtain

$$\begin{aligned} F(H,A) &= \sum_{B\subseteq A} (-1)^{|A|-|B|}\,\hat{Q}(H,B) \\ &= \sum_{B\subseteq A} (-1)^{|A|-|B|}\,Q(H-(U\setminus B)) \\ &= \sum_{U\setminus B\subseteq A} (-1)^{|A|-|U\setminus B|}\,Q(H-B) \\ &= (-1)^{|A|-|U|} \sum_{B\supseteq U\setminus A} (-1)^{|B|}\,Q(H-B). \end{aligned}$$

Similarly, we can prove for each $A \subseteq U$ that

$$F(K,A) = (-1)^{|A|-|U|} \sum_{B\supseteq U\setminus A} (-1)^{|B|}\,Q(K-B).$$

The substitution of $F(H,A)$ and $F(K,A)$ in equation (10) yields

$$\begin{aligned} Q(G) &= Q(H-U)\,Q(K-U) \\ &\quad + \frac{1}{y} \sum_{\emptyset\neq A\subseteq U} \frac{1}{x^{|A|}} \sum_{B\supseteq U\setminus A} (-1)^{|B|}\,Q(H-B) \sum_{C\supseteq U\setminus A} (-1)^{|C|}\,Q(K-C). \end{aligned}$$

□

Problem 26. *Can we have an analogue of Theorem 25 for the case where the separating vertex set is not required to be a clique?*

7. Reconstruction

The famous *graph reconstruction conjecture* by Kelly and Ulam [44] states the every undirected graph with at least three vertices can be reconstructed from a deck of its vertex-deleted subgraphs (more precisely from the corresponding isomorphism classes). See for example the papers [12, 11] as an introduction into this field. Despite the fact that the conjecture is still open, many graph invariants and graph polynomials (e.g. the Tutte polynomial) are known to be reconstructible. We can show that also the subgraph component polynomial can be reconstructed from the deck of the subgraph polynomials of the vertex-deleted subgraphs.

Theorem 27. *The subgraph polynomial $Q(G;x,y)$ for a graph $G=(V,E)$ with $n=|V(G)|\geq 3$ is uniquely determined by the set of polynomials*

$$\{Q(G-v;x,y) : v\in V(G)\}.$$

Let $\hat{\omega}$ be the smallest power of y that appears at least twice among the terms $x^{n-1}y^j$ of the polynomials $Q(G-v;x,y)$. Define

$$\omega = \begin{cases} n & \text{if } \hat{\omega} = n-1, \\ \hat{\omega} & \text{else.} \end{cases}$$

Then the subgraph polynomial is given by

$$Q(G;x,y) = x^n\left[y^{\omega} + \int_0^{\frac{1}{x}} t^{n-1} \sum_{v\in V} Q\left(G-v;\frac{1}{t},y\right)dt\right].$$

Proof. Let $k(G)$ denote the number of components of a graph G. In each graph $G=(V,E)$ with at least three vertices and at least one edge there exist two vertices $u,v \in V$ such that $k(G-u)=k(G-v)=k(G)$. If the term $x^{n-1}y^j$ appears in the polynomial $Q(G-v;x,y)$ then the number of components of $G-v$ equals j. Since $k(G-v) \geq k(G)$ for each vertex $v \in V$, the the smallest power of y that appears at least twice among the terms $x^{n-1}y^j$ of the polynomials $Q(G-v;x,y)$ is equal to $k(G)$. There is only one exception: If G is the empty (edgeless) graph then the removal of each vertex of G decreases the number of components by one, which is taken into consideration within the definition of ω. Consequently, we obtain $\omega = k(G)$.

Each vertex induced subgraph with $i < n$ vertices is counted exactly $n-i$ times in the polynomial

$$\sum_{v \in V(G)} Q(G-v;x,y).$$

The coefficient of $t^i y^j$ in

$$t^{n-1} \sum_{v \in V} Q\left(G-v;\frac{1}{t},y\right)$$

equals i times the number of vertex induced subgraphs of G with exactly $n-i$ vertices and j components. The integration with respect to t transforms t^{i-1} into $\frac{1}{i}t^i$ such that the vertex induced subgraphs are enumerated correctly by the coefficients of the resulting polynomial. Finally, the bounds of integration and the multiplication with x^n performs the back-substitution in order to obtain an ordinary generating function with variables x and y. □

8. Random Subgraphs

Now we assume that the vertices of $G=(V,E)$ fail stochastic independently with a given (identical) probability $q=1-p$. We obtain the probability $P_k(G)$ that a vertex induced subgraph of G has exactly k components from the subgraph polynomial:

$$P_k(G) = \frac{1}{k!}\frac{\partial^k}{\partial y^k}(1-p)^n Q\left(G;\frac{p}{1-p},y\right)\bigg|_{y=0} \tag{11}$$

The sequence $\{P_k(G)\}_{k \in \mathbb{N}}$ is the distribution of the number of components. Consequently, we obtain

$$\sum_{k=0}^{n} P_k(G) = 1.$$

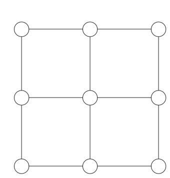

Figure 8. A 3×3 grid graph

Figure 9 shows the distribution for the graph presented in Figure 8.

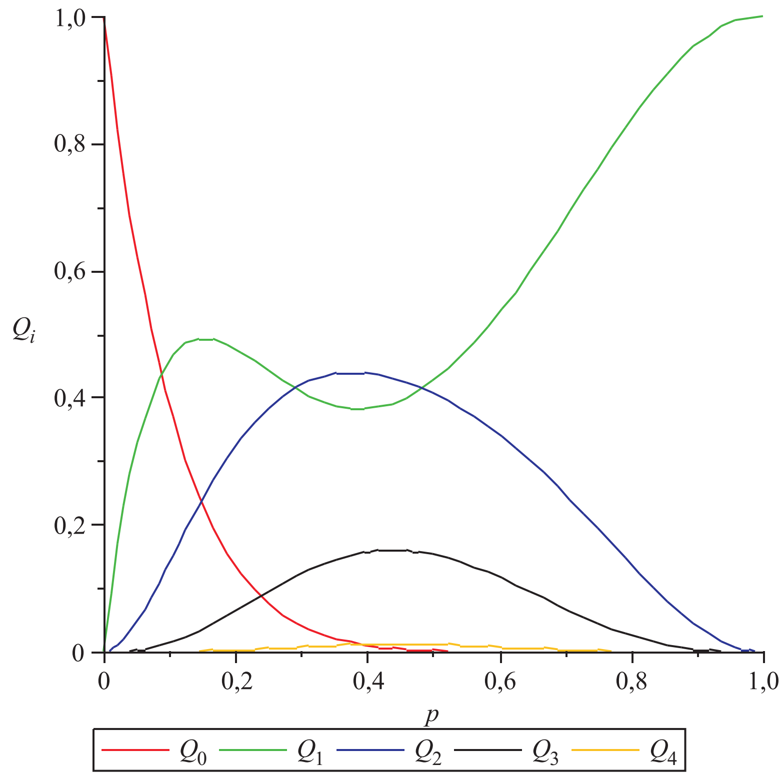

FIGURE 9. Distribution of the number of components

The probability $P_1(G)$ is called the *residual connectedness reliability.* Boesch, Satyanarayana, and Suffel [9] showed that the computation of $P_1(G)$ is a #P-hard problem, even in planar bipartite graphs. Since $P_1(G)$ can be obtained in polynomial time from the subgraph polynomial by applying the relation (11), we obtain the following statement.

Corollary 28. *The computation of the subgraph polynomial is a #P-hard problem. It remains #P-hard for the class of all planar bipartite graphs.*

9. COMPUTATIONAL COMPLEXITY OF $Q(G;x,y)$

9.1. **Complexity of evaluation.** We have already seen in Corollary 28 that $Q(G;x,y)$ is $\sharp\mathbf{P}$-hard to compute. Now we deal with a problem of evaluation of $Q(-;x,y)$ at a given point $(x,y)\in\mathbb{Q}^2$ for arbitrary input graph G.

Theorem 29. *For every point $(x,y)\in\mathbb{Q}^2$, possibly except for the lines $xy=0$, $y=1$, $x=-1$ and $x=-2$, the evaluation of $Q(G;x,y)$ for an input graph G is $\sharp\mathbf{P}$-hard.*

C. Hoffmann in [26] showed the following:

Theorem 30 (Hoffmann 2008)**.** *For every point $(x,y,z)\in\mathbb{Q}^3$, except possibly for the subsets $x=0$, $z=-xy$, $(x,z)\in\{(1,0),(2,0)\}$ and $y\in\{-2,-1,0\}$, the evaluation of $\xi(-;x,y,z)$ for an input graph G is $\sharp\mathbf{P}$-hard.*

Proof of Theorem 29: We use Theorem 30 and our Theorem (22). Under the conditions of Theorem (22), Hoffmann's exception sets are mapped to the lines $xy = 0$, $y = 1$, $x = -1$ and $x = -2$. It follows that for every point $(x, y) \in \mathbb{Q}^2$ that does not lay on one of those lines, the polynomial $Q(-; x, y)$ is $\sharp\mathbf{P}$-hard to evaluate even for an input line graph $L(G)$. □

The evaluation of $Q(-; x, y)$ is polynomial time computable for $xy = 0$ and for $y = 1$. It remains open whether it is polynomial time computable for $x = -1$ and $x = -2$. One can also ask, whether there is some point $(x, y) \in \mathbb{Q}^2$, in which $Q(-; x, y)$ is hard to evaluate for general input graph, but easy for input line graph.

9.2. **Parameterized complexity.** Here we discuss the computational complexity of $Q(G; x, y)$ for input graphs of bounded tree with, and for input graphs for bounded clique width. We do not need the exact definitions here. For background on tree-width the reader can consult [17]. Clique-width was defined in [15]. Both are discussed in [28].

Recall that the subgraph component polynomial is definable using the **MSOL**-formalism (definition 9) with auxiliary order, while the result is order-independent. Hense, using a general theorem from [29, 28], we have

Proposition 31. *$Q(G; x, y)$ is polynomial time computable on graphs of tree-width at most k where the exponent of the run time is independent of k.*

Moreover, applying the result of Courcelle, Makowsky and Rotics [14], combined with the results from [40], we have a similar result for graphs of bouded clique width:

Proposition 32. *$Q(G; x, y)$ is polynomial time computable on graphs of clique-width at most k where the exponent of the run time is independent of k.*

The drawback of the general methods of [29, 28] and [14], lies in the huge hidden constants, which make it practically unusable. However, an explicit dynamic algorithm for computing the polynomial $Q(G; x, y)$ on graphs of bounded tree-width, given the tree decomposition of the graph, where the constants are simply exponential in k, can be constructed along the same ideas as presented in [43, 22]. For the graphs of bounded clique width, given the clique decomposition of the graph, we know an algorithm with constants doubly-exponential in k. It is open whether an algorithm with constants simply exponential in k exists. For a comparison of the complexity of computing graph polynomials on graphs classes of bounded clique-width, cf. [33].

10. Conclusions and Open Problems

We have shown that $Q(G; x, y)$ is a universal vertex elimination polynomial. We have given a few combinatorial interpretations of its evaluations and coefficients. We have proven various splitting formulas for $Q(G; x, y)$ such as the multiplicativity, Theorem 13 and Theorem 25. Problem 26 asks for more such theorems. Besides having algorithmic importance, such splitting formulas increase our structural understanding of the graph polynomial under study, and may help us in analizing its distictive power.

We have looked at the graph polynomial $Q(G; x, y)$ from various angles and compared its behaviour and distinguishing power with the characteristic polynomial, the matching polynomial the Tutte polynomial and the universal edge elimination polynomial. We have not discussed the relationship of $Q(G;x,y)$ to other graph polynomials, such as the interlace polynomial, [3, 1], or the many other graph polynomials listed in [31].

We have seen that $Q(G;x,y)$ distinguishes between graphs where these polynomials do not. We have not found cases where these other polynomials do distinguish between graphs where $Q(G;x,y)$ does not. This is probably due to our lack of computerized tools for searching for such cases, cf. Problem 6. In Problem 23 we ask about comparing distinguishing power of $Q(G;x,y)$ and the universal edge elimination polynomial $\xi(G;x,y,z)$. This seems to be more tricky. We have given a few examples of graphs and graph families which are determined by $Q(G;x,y)$.

Problem 33. *Find more graph invariants which are determined by* $Q(G;x,y)$ *.*

Problem 34. *Find more classes of graphs which are determined by* $Q(G;x,y)$*.*

Returning to our motivation, we have only studied the simplest case of community structure in networks. We have studied the generating function of induced subgraphs with i vertices which have j components. More generally, one would want to study community structures where components are replaced by maximal k-connected components.

Problem 35. *What are the appropriate generating functions which capture the essence of various community structures?*

References

[1] M. Aigner and H. van der Holst. Interlace polynomials. *Linear Algebra and Applications*, 377:11–30, 2004.

[2] A. Andrzejak. Splitting formulas for Tutte polynomials. *Journal of Combinatorial Theory, Series B*, 70.2:346–366, 1997.

[3] R. Arratia, B. Bollobás, and G.B. Sorkin. The interlace polynomial of a graph. *Journal of Combinatorial Theory, Series B*, 92:199–233, 2004.

[4] I. Averbouch, B. Godlin, and J.A. Makowsky. The most general edge elimination polynomial. arXiv `http://uk.arxiv.org/pdf/0712.3112.pdf`, 2007.

[5] I. Averbouch, B. Godlin, and J.A. Makowsky. An extension of the bivariate chromatic polynomial. submitted, 2008.

[6] M. Bläser and H. Dell. Complexity of the cover polynomial. In L. Arge, C. Cachin, T. Jurdziński, and A. Tarlecki, editors, *Automata, Languages and Programming, ICALP 2007*, volume 4596 of *Lecture Notes in Computer Science*, pages 801–812. Springer, 2007.

[7] M. Bläser and H. Dell. Complexity of the Bollobás-Riordan polynomia. exceptional points and uniform reductions. In Edward A. Hirsch, Alexander A. Razborov, Alexei Semenov, and Anatol Slissenko, editors, *Computer Science–Theory and Applications, Third International Computer Science Symposium in Russia*, volume 5010 of *Lecture Notes in Computer Science*, pages 86–98. Springer, 2008.

[8] Markus Bläser and Christian Hoffmann. On the complexity of the interlace polynomial. In *STACS*, pages 97–108, 2008.

[9] F. Boesch, A. Satyanarayana, and C. Suffel. On residual connectedness reliability. In F. Roberts, F. Hwang, and C. Monma, editors, *Reliability of Computer and Communication Networks*, volume 5 of *DIMACS Series in Discrete Math. and Theor. Comp. Science*, pages 51–59. AMS and ACM, 1991.

[10] B. Bollobás. *Modern Graph Theory*. Springer, 1999.

[11] J. A. Bondy. A graph reconstructor's manual. In A.D. Keedwell, editor, *Surveys in Combinatorics, 1991*, volume 166 of *London Mathematical Society Lecture Note Series*, pages 221–252. Cambridge University Press, 1991.

[12] J. A. Bondy and R.L. Hemminger. Graph reconstruction: A survey. *Journal of Graph Theory*, 1:227–268, 1977.

[13] B. Courcelle. *Graph Structure and Monadic Second Order Logic*. Cambridge University Press, in preparation.
[14] B. Courcelle, J.A. Makowsky, and U. Rotics. On the fixed parameter complexity of graph enumeration problems definable in monadic second order logic. *Discrete Applied Mathematics*, 108(1-2):23–52, 2001.
[15] B. Courcelle and S. Olariu. Upper bounds to the clique–width of graphs. *Discrete Applied Mathematics*, 101:77–114, 2000.
[16] A. de Mier and M. Noy. On graphs determined by their Tutte polynomials. *Graphs and Combinatorics*, 20.1:105–119, 2004.
[17] R. Diestel. *Graph Decompositions, A Study in Infinite Graph Theory*. Clarendon Press, Oxford, 1990.
[18] K. Dohmen, A. Pönitz, and P. Tittmann. A new two-variable generalization of the chromatic polynomial. *Discrete Mathematics and Theoretical Computer Science*, 6:69–90, 2003.
[19] R.G. Downey and M.F Fellows. *Parametrized Complexity*. Springer, 1999.
[20] M. Dyer and C. Greenhill. The complexity of counting graph homomorphisms. *Random Structures and Algorithms*, 17.3-4:260 – 289, 2000.
[21] H. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer Verlag, 1995.
[22] E. Fischer, J.A. Makowsky, and E.V. Ravve. Counting truth assignments of formulas of bounded tree width and clique-width. *Discrete Applied Mathematics*, 156:511–529, 2008.
[23] M. Freedman, László Lovász, and A. Schrijver. Reflection positivity, rank connectivity, and homomorphisms of graphs. *Journal of AMS*, 20:37–51, 2007.
[24] M. Girvan and M.E.J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, 99:7821–7826, 2002.
[25] B. Godlin, E. Katz, and J.A. Makowsky. Graph polynomials: From recursive definitions to subset expansion formulas. arXiv `http://uk.arxiv.org/pdf/0812.1364.pdf`, 2008.
[26] C. Hoffmann. A most general edge elimination polynomial–thickening of edges. arXiv:0801.1600v1 [math.CO], 2008.
[27] F. Jaeger, D.L. Vertigan, and D.J.A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Math. Proc. Camb. Phil. Soc.*, 108:35–53, 1990.
[28] J.A. Makowsky. Algorithmic uses of the Feferman-Vaught theorem. *Annals of Pure and Applied Logic*, 126.1-3:159–213, 2004.
[29] J.A. Makowsky. Colored Tutte polynomials and Kauffman brackets on graphs of bounded tree width. *Disc. Appl. Math.*, 145(2):276–290, 2005.
[30] J.A. Makowsky. From a zoo to a zoology: Descriptive complexity for graph polynomials. In A. Beckmann, U. Berger, B. Löwe, and J.V. Tucker, editors, *Logical Approaches to Computational Barriers, Second Conference on Computability in Europe, CiE 2006, Swansea, UK, July 2006*, volume 3988 of *Lecture Notes in Computer Science*, pages 330–341. Springer, 2006.
[31] J.A. Makowsky. From a zoo to a zoology: Towards a general theory of graph polynomials. *Theory of Computing Systems*, 43:542–562, 2008.
[32] J.A. Makowsky and E. Fischer. Linear recurrence relations for graph polynomials. In A. Avron, N. Dershowitz, and A. Rabinowitz, editors, *Boris (Boaz) A. Trakhtenbrot on the occasion of his 85th birthday*, volume 4800 of *LNCS*, pages 266–279. Springer, 2008.
[33] J.A. Makowsky, U. Rotics, I. Averbouch, and B. Godlin. Computing graph polynomials on graphs of bounded clique-width. In F. V. Fomin, editor, *Graph-Theoretic Concepts in Computer Science, 32nd International Workshop, WG 2006, Bergen, Norway, June 22-23, 2006, Revised Papers*, volume 4271 of *Lecture Notes in Computer Science*, pages 191–204. Springer, 2006.
[34] J. Nesetril and P. Winkler, editors. *Graphs, Morphisms and Statistical Physics*, volume 63 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. AMS, 2004.
[35] M.E.J. Newman. Detecting community structure in networks. *Eur. Phys. J.B.*, 38:321–330, 2004.
[36] M.E.J. Newman, A.L. Barabasi, and D. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, 2006.
[37] S.D. Noble. Evaluating the Tutte polynomial for graphs of bounded tree-width. *Combinatorics, Probability and Computing*, 7:307–321, 1998.
[38] M. Noy. On graphs determined by polynomial invariants. *tcs*, 307:365–384, 2003.

[39] M. Noy and A. Ribó. Recursively constructible families of graphs. *Advances in Applied Mathematics*, 32:350–363, 2004.

[40] S. Oum. Approximating rank-width and clique-width quickly. In *Graph Theoretic Concepts in Computer Science, WG 2005*, volume 3787 of *Lecture Notes in Computer Science*, pages 49–58, 2005.

[41] J.G. Oxley and D.J.A. Welsh. The Tutte polynomial and percolation. In J.A. Bundy and U.S.R. Murty, editors, *Graph Theory and Related Topics*, pages 329–339. Academic Press, London, 1979.

[42] L. Traldi. A subset expansion of the coloured Tutte polynomial. *Combinatorics, Probability and Computing*, 13:269–275, 2004.

[43] L. Traldi. On the colored Tutte polynomial of a graph of bounded tree-width. *Discrete Applied Mathematics*, 154.6:1032–1036, 2006.

[44] S.M. Ulam. *A Collection of Mathematical Problems*. Wiley, 1960.

E-mail address, P. Tittmann: peter@HTWM.De
E-mail address, I. Averbouch: ailia@cs.technion.ac.il
E-mail address, J.A. Makowsky: janos@cs.technion.ac.il

(P. Tittmann) Fachbereich Mathematik–Physik–Informatik, Hochschule Mittweida, Mittweida, Germany

(I. Averbouch and J.A. Makowsky) Department of Computer Science, Technion–Israel Institute of Technology, 32000 Haifa, Israel