

Connected components and evolution of random graphs: an algebraic approach

René Schott · G. Stacey Staples

Received: 26 June 2009 / Accepted: 19 May 2011 / Published online: 17 June 2011
© Springer Science+Business Media, LLC 2011

Abstract Questions about a graph’s connected components are answered by studying appropriate powers of a special “adjacency matrix” constructed with entries in a commutative algebra whose generators are idempotent. The approach is then applied to the Erdős–Rényi model of sequences of random graphs. Developed herein is a method of encoding the relevant information from graph processes into a “second quantization” operator and using tools of quantum probability and infinite-dimensional analysis to derive formulas that reveal the exact values of quantities that otherwise can only be approximated. In particular, the expected size of a maximal connected component, the probability of existence of a component of particular size, and the expected number of spanning trees in a random graph are obtained.

Keywords Random graphs · Graph processes · Quantum probability

1 Introduction

The evolution of random graphs has been studied in some detail. The first works in this area are attributed to Erdős and Rényi [9–11].

Definition 1.1 Let n be a positive integer, let $V = \{1, 2, \dots, n\}$, and define $\Omega = \binom{V}{2}$. A *graph process* on V is a sequence $(G_t)_{t=0}^{\Omega} = (V, E_t)_{t=0}^{\Omega}$ such that each G_t is a graph on V with t edges, and $G_0 \subset G_1 \subset \dots \subset G_{\Omega}$.

R. Schott
IECN and LORIA Nancy Université, Université Henri Poincaré, BP 239,
54506 Vandoeuvre-lès-Nancy, France
e-mail: schott@loria.fr

G.S. Staples (✉)
Department of Mathematics and Statistics, Southern Illinois University Edwardsville, Edwardsville,
IL 62026-1653, USA
e-mail: [sstaple@siue.edu](mailto:ssstaple@siue.edu)

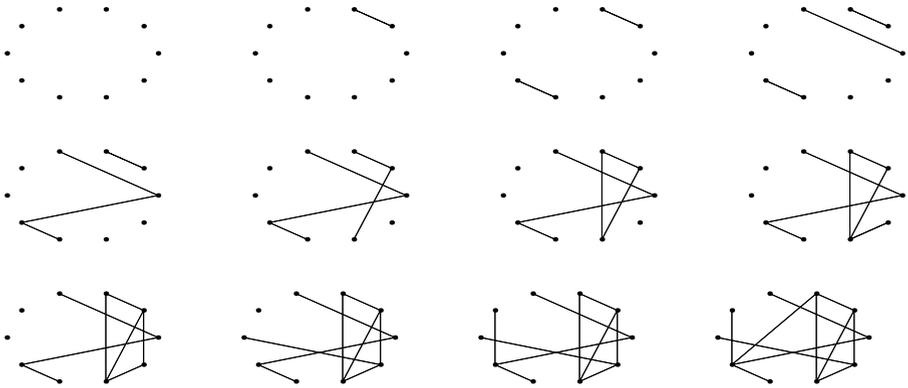


Fig. 1 A graph process on 10 vertices

Let $\hat{\mathcal{G}}$ denote the probability space formed by the set of all $\Omega!$ graph processes with equal probability defined for all. It is well known that any graph process is a Markov chain whose states are graphs on V .

Example 1.2 Graphs G_0 through G_{11} of a graph process on ten vertices are pictured in Fig. 1. Note that G_{11} is the first connected graph of the sequence.

Erdős and Rényi proved that if $t \sim cn$ for some fixed $c \in \mathbb{R}$ where $0 < c < \frac{1}{2}$, then almost every G_t is such that its largest component has $O(\log n)$ vertices. If $c > \frac{1}{2}$, then the largest component of almost every G_t has $(1 - \alpha_c + o(1))n$ vertices for some $0 < \alpha_c < 1$. Finally, if $t = \lfloor n/2 \rfloor$, then the maximal size of a component of almost every G_t is $O(n^{2/3})$.

Another notable work is that of Bollobás [7], who showed that almost every $G \in \hat{\mathcal{G}}$ is such that for $t \geq n/2 + (\log n)^{1/2}n^{2/3}$, the graph G_t has a unique component of order at least $n^{2/3}$, referred to as the *giant component*.

Considering “online” processes, Bohman and Frieze [3] investigated algorithms for avoiding the emergence of the giant component in a graph process. Bohman, Frieze, and Wormald [4] and Bohman and Kim [5] also considered avoiding the giant component.

In contrast, Flaxman, Gamarnik, and Sorking [12] and Bohman and Kravitz [6] considered algorithms for obtaining a giant component in a graph process.

Chung and Lu [8] investigated the distribution of the sizes of the connected components in a family of random graphs with given expected degree sequence.

Frieze and Łuczak [13] considered maximal numbers of edge-disjoint spanning trees in random graphs. In related work, Palmer and Spencer [17] showed that in almost every random graph process, the hitting time for having k edge-disjoint spanning trees equals the hitting time for having minimum degree k .

Also of interest is the investigation by Molloy [15] of the connections between satisfiability thresholds for random k -SAT and thresholds for the emergence of the giant component in a graph process.

What is proposed in this paper is an algebraic framework in which many quantities related to a random graph’s connected components can be expressed explicitly.

Defined here is an adjacency matrix whose entries lie in a commutative algebra with idempotent generators. After labeling the graph’s edges with idempotent generators of the algebra, computing powers of this matrix reveals information about the graph’s connected components.

The idempotent-adjacency matrix approach is then extended to graph processes by creating a *second quantization* space of graph processes. All possible graph processes are encoded in one operator, and information about the connected components contained in the N th graph of the sequence is revealed by considering powers of this operator.

The method of second quantization is well known to physicists and goes back to the original work of Berezin [2].

Quantum probabilistic approaches to graph theory are also not new. Hashimoto, Hora, and Obata [14] used the method of quantum decomposition to obtain central limit theorems for growing sequences of graphs. Other applications of quantum probabilistic techniques to graph theory include the work of Obata [16] and Accardi et al. [1].

Historically, the graph-theoretic work done by quantum probabilists has dealt with specific graphs whose relationship to problems of mathematical physics are understood. In contrast, the philosophy of the current authors is that the tools of quantum probability can be applied to more general graph-theoretic problems.

In earlier work, the current authors defined nilpotent adjacency matrices and applied them to the study of cycles in random graphs [19, 20]. A similar approach based on a commutative algebra whose generators $\{\zeta_i\}$ satisfy $\zeta_i^2 = 1$ has been used to formulate random walks on the hypercube [21].

1.1 Algebraic preliminaries

Throughout this paper, let \mathbb{N} denote the set of *positive* integers. The notation \mathbb{N}_0 will be used to denote the set $\mathbb{N} \cup \{0\}$.

Definition 1.3 Let V be a finite set with $n > 0$ elements. Let \mathcal{I}_V be the associative algebra over \mathbb{R} generated by commuting idempotents $\{\varepsilon_{\{i\}} : i \in V\}$ along with the unit scalar $\varepsilon_{\emptyset} = 1 \in \mathbb{R}$. In particular, for $i, j \in V$, the generators satisfy

$$\varepsilon_{\{i\}}\varepsilon_{\{j\}} = \varepsilon_{\{j\}}\varepsilon_{\{i\}} \tag{1.1}$$

and

$$\varepsilon_{\{i\}}\varepsilon_{\{i\}} = \varepsilon_{\{i\}}. \tag{1.2}$$

For simplicity of notation, linear basis elements will be indexed by subsets of the power set 2^V ; i.e.,

$$\varepsilon_{\underline{l}} = \prod_{i \in \underline{l}} \varepsilon_i. \tag{1.3}$$

Remark 1.4 An easy realization is the algebra generated by canonical projections onto orthogonal hyperplanes in \mathbb{R}^n . For example, consider the collection $\{\varepsilon_i\}_{1 \leq i \leq n}$ defined by

$$\varepsilon_i(x_1, \dots, x_n) = (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n). \tag{1.4}$$

Note that any element $x \in \mathcal{I}_V$ has canonical expansion of the form

$$x = \sum_{\underline{i} \in 2^V} x_{\underline{i}} \varepsilon_{\underline{i}}, \tag{1.5}$$

where $x_{\underline{i}} \in \mathbb{R}$ for each $\underline{i} \in 2^V$.

Define the *grade* of $x \in \mathcal{I}_V$ by the mapping $\lambda : \mathcal{I}_V \rightarrow \mathbb{N}_0$ satisfying

$$\lambda(x) = \max_{x_{\underline{i}} \neq 0} |\underline{i}|. \tag{1.6}$$

Hence, the grade of x is the size of a maximal multi-index in the canonical expansion of x . For example, $\lambda(1 + \varepsilon_{\{1,2\}} + 3\varepsilon_{\{1,3,4\}} - 4\varepsilon_{\{1,2,3,4,5\}}) = 5$.

Remark 1.5 Letting $n = |V|$, it is worth noting that the idempotent-generated algebra \mathcal{I}_V can be constructed within the $2n$ -particle fermion creator/annihilator algebra [19].

Definition 1.6 Let $\{\varepsilon_{\{i\}} : i \in V\}$ denote the idempotent generators of \mathcal{I}_V . Associated with any finite graph $G = (V, E)$ on n vertices is a *column idempotent-adjacency matrix* a defined by

$$a_{ij} = \begin{cases} \varepsilon_{\{j\}} & \text{if } i = j, \\ \varepsilon_{\{j\}} & \text{if } (v_i, v_j) \in E \subset V \times V, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

Further, define the notation a^\dagger to be the matrix transpose of a .

Defining the diagonal matrix Δ by $\Delta_{ii} = \varepsilon_{\{i\}}$, the column idempotent-adjacency matrix a associated with a finite graph satisfies

$$a = (A + I)\Delta, \tag{1.8}$$

where A denotes the usual adjacency matrix of the graph. The addition of the identity matrix I is necessary to account for any isolated vertices in the graph.

For any column idempotent-adjacency matrix a , the transpose a^\dagger satisfies

$$a^\dagger_{ij} = \begin{cases} \varepsilon_{\{i\}} & \text{if } i = j, \\ \varepsilon_{\{i\}} & \text{if } (v_i, v_j) \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.9}$$

Moreover,

$$a^\dagger = \Delta(A + I)^\dagger = \Delta(A + I), \tag{1.10}$$

where A denotes the usual adjacency matrix of the graph. For this reason, a^\dagger will be referred to as a *row idempotent-adjacency matrix*.

For each positive integer n , the collection of $n \times n$ matrices over \mathcal{I}_V constitutes a unital $*$ -algebra \mathcal{A} with the usual matrix identity and involution $a^* = a^\dagger$. Within this $*$ -algebra, the column idempotent-adjacency matrix $(A + I)\Delta$ generates a multiplicative semigroup with right identity Δ . This semigroup is defined by

$$\mathcal{G}^c = \{((A + I)\Delta)^\ell : \ell \in \mathbb{N}\}. \tag{1.11}$$

Likewise, within this $*$ -algebra, the row idempotent-adjacency matrix $\Delta(A + I)$ generates a multiplicative semigroup with left identity Δ . This semigroup is defined by

$$\mathcal{G}^r = \{(\Delta(A + I))^\ell : \ell \in \mathbb{N}\}. \tag{1.12}$$

The results in the remainder of the paper hold for either choice of idempotent-adjacency matrix. Without loss of generality, fix $\mathcal{G} = \mathcal{G}^r$ or $\mathcal{G} = \mathcal{G}^c$ throughout the remainder of the paper.

Letting $\{\mathbf{e}_i\}_{1 \leq i \leq |V|}$ be the standard basis for $\mathbb{R}^{|V|}$ taken as column vectors, a linear mapping $\mathcal{G} \rightarrow \mathcal{I}_V^{|V|}$ is naturally induced for each i by $a \mapsto a \mathbf{e}_i$. Using Dirac notation, $\langle \mathbf{e}_i | a \mathbf{e}_i \rangle := \mathbf{e}_i^\dagger a \mathbf{e}_i = a_{ii}$. Moreover, define the notation $\rho_i := |\mathbf{e}_i\rangle\langle \mathbf{e}_i|$.

The *trace* of $a \in \mathcal{G}$ is the linear mapping $\tau : \mathcal{G} \rightarrow \mathcal{I}_V$ defined by

$$\tau(a) := \sum_{i=1}^{|V|} \langle \mathbf{e}_i | a \mathbf{e}_i \rangle. \tag{1.13}$$

The mapping τ will also be considered a state on \mathcal{G} , making the pair (\mathcal{G}, τ) an *algebraic probability space*.

For positive integer m , we refer to $\langle \mathbf{e}_i | a^m \mathbf{e}_i \rangle = \tau(\rho_i a^m)$ as the *mth moment of a in the state \mathbf{e}_i* . Further, we refer to $\tau(a^m)$ as the *mth moment of a in the state τ* .

2 Connected components

Proposition 2.1 *Let G be a simple graph on n vertices $\{v_1, \dots, v_n\}$, let a denote an idempotent-adjacency matrix for G , and let C_i ($1 \leq i \leq n$) denote the size of the maximal connected component of G containing vertex v_i . Then,*

$$\lambda(\tau(\rho_i a^{n(n-1)})) = C_i. \tag{2.1}$$

Proof A straightforward inductive argument shows that for any positive integer k the matrix entry $(a^k)_{ij}$ corresponds to the collection of k -walks from the i th vertex

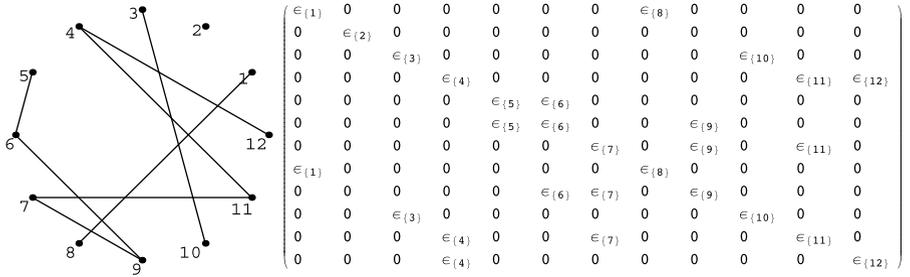


Fig. 2 A simple graph on 12 vertices and its idempotent-adjacency matrix

to the j th vertex in the graph. By definition of the adjacency matrix, all vertices included in such a walk are contained in the same connected component of the graph. By construction of the idempotent-adjacency matrix, the grade of $(a^k)_{ij}$ reveals the maximum number of *distinct* vertices contained in any k -walk from v_i to v_j . In the worst case, a closed walk repeats every edge in covering the vertices of a connected component, hence $2\binom{n}{2} = n(n - 1)$ steps are allowed. □

By Proposition 2.1, an idempotent-adjacency matrix of a graph on n vertices is a quantum random variable whose $(n^2 - n)$ th moment in the state e_i reveals the size of the maximal component containing the graph’s i th vertex. As a corollary, summing the reciprocals of the grades of n th moments over the states e_i gives the number of connected components in the graph.

Corollary 2.2 *Let G be a simple graph on n vertices, let a denote an associated idempotent-adjacency matrix, and let C denote the number of connected components of G . Then,*

$$\sum_{i=1}^n \frac{1}{\lambda(\tau(\rho_i a^{n(n-1)}))} = C. \tag{2.2}$$

Proof Note that by construction of the idempotent-adjacency matrix, $\lambda(\tau(\rho_i a^{n(n-1)})) \geq 1$ for all $1 \leq i \leq n$. The result then follows from Proposition 2.1. □

Example 2.3 Consider the graph of Fig. 2. Direct computation shows the grade of the trace to be 7, corresponding to the maximal connected component consisting of vertex set $\{v_4, v_5, v_6, v_7, v_9, v_{11}, v_{12}\}$.

Examining the diagonal entries of the 22nd power of the matrix appearing in Fig. 2, reveals components (and sub-components) containing each vertex.

Vertex	Maximum grade term
v_1	$2097151 \varepsilon_{\{1,8\}}$
v_2	$\varepsilon_{\{2\}}$
v_3	$2097151 \varepsilon_{\{3,10\}}$
v_4	$28633021 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_5	$13224244 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_6	$28633021 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_7	$31249000 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_8	$2097151 \varepsilon_{\{1,8\}}$
v_9	$31033277 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_{10}	$2097151 \varepsilon_{\{3,10\}}$
v_{11}	$31033277 \varepsilon_{\{4,5,6,7,9,11,12\}}$
v_{12}	$13224244 \varepsilon_{\{4,5,6,7,9,11,12\}}$

Proposition 2.4 *Let G be a simple graph on n vertices, let a denote an associated idempotent-adjacency matrix, and let M denote the size of a maximal connected component in G . Then,*

$$\lambda(\tau(a^{n(n-1)})) = M. \tag{2.3}$$

Proof An inductive argument shows that the diagonal entries of a^{n^2-n} are sums of idempotents representing closed walks of length $n^2 - n$ on the graph. Because the graph contains n vertices, the maximal connected component of G will be covered by a closed walk of length $n(n - 1)$ or less. All components can be covered by closed walks of length equal to $n^2 - n$ by the inclusion of a loop based at each vertex in the definition of the idempotent-adjacency matrix. □

By Proposition 2.4, an idempotent-adjacency matrix of a graph on n vertices is a quantum random variable whose $(n^2 - n)$ th moment in the state τ corresponds to the graph’s connected components. The corresponding grade is the maximum size among the connected components.

Usually, the unique largest component of a graph is called the *giant component* [7]. In general, one can define a giant component as one containing a fixed proportion of the graph’s vertices.

Noting that in a graph on n vertices, any connected component containing a majority of the graph’s vertices is the unique largest component, we define the giant component accordingly.

Definition 2.5 Given a simple graph G on n vertices, the *giant component* of G , if it exists, is defined as the unique connected component of order greater than $n/2$. Equivalently, any component containing a majority of the vertices of G is the giant component of G .

Corollary 2.6 *A simple graph G on n vertices contains a giant component if and only if $\lambda(\tau(a^{n^2-n})) > \frac{n}{2}$, where a is an idempotent-adjacency matrix of the graph.*

2.1 (k, d) -components

Definition 2.7 A component of a graph is said to be a (k, d) -component if it has k vertices and $k + d$ edges.

When considering (k, d) -components, it quickly becomes apparent that the existing idempotent-adjacency matrix construction is inadequate. To address this, it is necessary to label edges as well as vertices with idempotents.

Definition 2.8 Let V be a finite set with $n > 0$ elements. Let $\mathcal{I}_{V \times V}$ be the associative algebra over \mathbb{R} generated by commuting idempotents

$$\{\gamma_{\{(i,j)\}} : (i, j) \in V \times V\}$$

along with the unit scalar $\gamma_\emptyset = 1 \in \mathbb{R}$.

In particular, for $(i, j), (k, \ell) \in V \times V$, the generators of $\mathcal{I}_{V \times V}$ satisfy

$$\gamma_{\{(i,j)\}}\gamma_{\{(k,\ell)\}} = \gamma_{\{(k,\ell)\}}\gamma_{\{(i,j)\}} \tag{2.4}$$

and

$$\gamma_{\{(i,j)\}}\gamma_{\{(i,j)\}} = \gamma_{\{(i,j)\}}. \tag{2.5}$$

To simplify notation, linear basis elements will be indexed by subsets of the power set $2^{V \times V}$; i.e.,

$$\gamma_{\underline{i}} = \prod_{(i,j) \in \underline{i}} \gamma_{\{(i,j)\}}. \tag{2.6}$$

Definition 2.9 Let $\{\varepsilon_{\{i\}} : i \in V\}$ denote the idempotent generators of \mathcal{I}_V , and let $\{\gamma_{\{(i,j)\}} : i, j \in V\}$ denote the idempotent generators of $\mathcal{I}_{V \times V}$. Associated with any finite graph $G = (V, E)$ on n vertices is a *vertex/edge-labeled idempotent-adjacency matrix* \hat{a} having entries in $\mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ defined by

$$\hat{a}_{ij} = \begin{cases} \varepsilon_{\{j\}} \otimes 1 & \text{if } i = j, \\ \varepsilon_{\{j\}} \otimes \gamma_{\{(i,j)\}} & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases} \tag{2.7}$$

Further, for nonnegative integers m and ℓ , the m, ℓ -grade projection is defined by

$$\left\langle \sum_{\underline{i} \in 2^V} \sum_{\underline{j} \in 2^{V \times V}} \alpha_{\underline{i}, \underline{j}} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}} \right\rangle_{m, \ell} = \sum_{|\underline{i}|=m} \sum_{|\underline{j}|=\ell} \alpha_{\underline{i}, \underline{j}} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}. \tag{2.8}$$

Given arbitrary element $u \in \mathcal{I}_V \otimes \mathcal{I}_{V \times V}$, the notation $\dim(u)$ will denote the dimension of the smallest subspace S of $\mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ such that $u \in S$.

Definition 2.10 A term $u_{\underline{i},j} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}$ of the canonical expansion of an element $u \in \mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ will be said to be a *top form* for u if for every term $u_{\underline{\ell},k} \varepsilon_{\underline{\ell}} \otimes \gamma_{\underline{k}}$ of the canonical expansion of u , the following conditions hold: (i) $|\underline{i}| \geq |\underline{\ell}|$, and (ii) $\underline{i} = \underline{\ell} \Rightarrow |\underline{j}| \geq |\underline{k}|$.

Definition 2.11 Define the *top-form projection* μ on $\mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ by

$$\mu(u) = \sum_{\text{top forms}} u_{\underline{i},j} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}. \tag{2.9}$$

Define \mathcal{G}_e to be the multiplicative semigroup generated by vertex/edge-labeled idempotent-adjacency matrices, and extend the trace mapping to $\tau : \mathcal{G}_e \rightarrow \mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ in the natural way; i.e., the trace of $a \in \mathcal{G}_e$ is the linear mapping $\tau : \mathcal{G}_e \rightarrow \mathcal{I}_V \otimes \mathcal{I}_{V \times V}$ defined by

$$\tau(a) := \sum_{i=1}^{|V|} \langle \mathbf{e}_i | a | \mathbf{e}_i \rangle. \tag{2.10}$$

Proposition 2.12 Let G be a simple graph on n vertices $\{v_1, \dots, v_n\}$, let \hat{a} denote a vertex/edge-labeled idempotent-adjacency matrix for G . Then, for fixed positive integers k and d , vertex v_i is contained in a (k, d) -component of G if and only if

$$\langle \tau(\rho_i \hat{a}^{n(n-1)}) \rangle_{k,k+d} = \mu(\tau(\rho_i \hat{a}^{n(n-1)})). \tag{2.11}$$

Proof As in Proposition 2.1, vertex v_i is contained in a maximal connected component on k vertices if and only if the top-form component of $\tau(\rho_i \hat{a}^{n(n-1)})$ is of the form $u_{\underline{i},j} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}$ with $|\underline{i}| = k$. This component is a (k, d) component if and only if $|\underline{j}| = k + d$. □

Proposition 2.12 says that the vertex/edge-labeled idempotent-adjacency matrix of a graph is a quantum random variable whose $(n^2 - n)$ th moment in the state \mathbf{e}_i corresponds to the vertices and edges in the (k, d) components containing vertex v_i .

Proposition 2.13 Let G be a simple graph on n vertices, let \hat{a} denote the associated vertex/edge-labeled idempotent-adjacency matrix, and let $C(k, d)$ denote the number of (k, d) -components of G . If $\langle \tau(\rho_i \hat{a}^{n(n-1)}) \rangle_{k,k+d} = \mu(\tau(\rho_i \hat{a}^{n(n-1)}))$, then

$$\dim(\mu(\tau(\hat{a}^{n(n-1)}))) = \sharp\{(k, d)\text{-components of } G\}. \tag{2.12}$$

Proof The edges and vertices of each (k, d) component are represented by a unique basis element $\varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}$. The result follows immediately from Proposition 2.12 by summing over states. □

In particular, when G is a simple graph on n vertices with associated vertex/edge-labeled idempotent-adjacency matrix \hat{a} , $\langle \tau(\rho_i \hat{a}^{n(n-1)}) \rangle_{k,k-1} = \mu(\tau(\rho_i \hat{a}^{n(n-1)}))$ implies

$$\dim(\langle \tau(\hat{a}^{n(n-1)}) \rangle_{k,k-1}) = \sharp\{k\text{-vertex tree components of } G\}. \tag{2.13}$$

Proposition 2.14 *Let \hat{a} denote the vertex/edge-labeled idempotent-adjacency matrix of a simple graph $G = (V, E)$ on n vertices. Then,*

$$\dim(\langle \tau(\hat{a}^{n^2-n}) \rangle_{n,n-1}) = \sharp\{\text{spanning trees of } G\}. \tag{2.14}$$

Proof By construction of the vertex/edge-labeled idempotent-adjacency matrix, nonzero terms of $\langle \tau(\hat{a}^{n^2-n}) \rangle_{n,n-1}$ correspond to connected components on n vertices and $n - 1$ edges, i.e., spanning trees. The subsets \underline{i} and \underline{j} indexing $\varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}$ in these terms specify the vertices and edges contained in the spanning tree, respectively. Distinct blades in $\mathcal{I}_{V \times V}$ correspond to distinct edge sets, and thus distinct spanning trees. □

Let $\{c_{(k,\ell)} : (k, \ell) \in V \times V\} \subset \mathbb{R}$ denote a collection of costs associated with the edges in $G = (V, E)$. Let \hat{a} denote the vertex/edge-labeled idempotent-adjacency matrix for G , and let \mathcal{B} denote the set of basis elements $\varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}$ for $\langle \tau(\hat{a}^{n^2-n}) \rangle_{n,n-1}$. Define the mapping $\varpi : \mathcal{B} \rightarrow \mathbb{R}$ by

$$\varpi(\varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}) = \prod_{(k,\ell) \in \underline{j}} e^{-c_{(k,\ell)}}. \tag{2.15}$$

Corollary 2.15 *Let G be a simple graph on n vertices, and let \hat{a} denote the associated vertex/edge-labeled idempotent-adjacency matrix. Then, a minimum cost spanning tree of G has cost C_T , given by*

$$C_T = -\ln\left(\max_{x \in \mathcal{B}} \{\varpi(x)\}\right). \tag{2.16}$$

The edge sets of the minimum cost spanning trees are determined by the corresponding elements of \mathcal{B} .

Definition 2.16 Let $\mathcal{N}_{V \times V}$ denote the associative algebra over \mathbb{R} generated by commuting null-squares

$$\{\zeta_{\{(i,j)\}} : (i, j) \in V \times V\}$$

along with the unit scalar $\zeta_{\emptyset} = 1 \in \mathbb{R}$. In particular, for $(i, j), (k, \ell) \in V \times V$, the generators of $\mathcal{N}_{V \times V}$ satisfy

$$\zeta_{\{(i,j)\}} \zeta_{\{(k,\ell)\}} = \zeta_{\{(k,\ell)\}} \zeta_{\{(i,j)\}} \tag{2.17}$$

and

$$\zeta_{\{(i,j)\}} \zeta_{\{(i,j)\}} = 0. \tag{2.18}$$

To simplify notation, linear basis elements will be indexed by subsets of the power set $2^{V \times V}$; i.e.,

$$\zeta_{\underline{i}} = \prod_{(i,j) \in \underline{i}} \zeta_{\{(i,j)\}}. \tag{2.19}$$

Remark 2.17 The authors have used algebras generated by commuting nilpotents to treat a number of problems related to enumerating cycles and self-avoiding walks in graphs (cf. [19, 20]).

Now define the mapping $\Psi : \mathcal{I}_V \otimes \mathcal{I}_{V \times V} \rightarrow \mathcal{I}_V \otimes \mathcal{N}_{V \times V}$ by linear extension of $\alpha \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}} \mapsto \alpha \varepsilon_{\underline{i}} \otimes \zeta_{\underline{j}}$, and let $\hat{\mathcal{B}}$ denote the sum of elements in \mathcal{B} . That is,

$$\hat{\mathcal{B}} = \sum_{\varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}} \in \mathcal{B}} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}}. \tag{2.20}$$

The nilpotent properties of $\mathcal{N}_{V \times V}$ now make it possible to sieve out pairwise edge-disjoint spanning trees.

Proposition 2.18 *Let G be a simple graph on n vertices, and let \hat{a} denote the associated vertex/edge-labeled idempotent-adjacency matrix. Let $\hat{\mathcal{B}}$ be defined as in (2.20). Let D_s denote the size of a maximal collection of pairwise edge-disjoint spanning trees of G . Then,*

$$D_s = \text{deg}_t \exp(t\Psi(\hat{\mathcal{B}})). \tag{2.21}$$

In other words, D_s is equal to the degree of $\exp(t\Psi(\hat{\mathcal{B}}))$ as a polynomial in t .

Proof The proposition is a corollary of Proposition 2.14. By the nilpotent properties of $\mathcal{N}_{V \times V}$, straightforward induction reveals that for each $\ell > 0$,

$$(t\Psi(\hat{\mathcal{B}}))^\ell = t^\ell \ell! \sum_{\substack{\text{pairwise-disjoint } \ell\text{-tuples} \\ \underline{\zeta}_{j_1}, \dots, \underline{\zeta}_{j_\ell}}} \prod_{k=1}^{\ell} \varepsilon_{i_k} \otimes \zeta_{j_k}. \tag{2.22}$$

Thus, the multivectors associated with pairwise edge-disjoint ℓ -tuples of spanning trees are recovered with multiplicity $\ell!$ from the ℓ th power. The multiplicity factor is removed by considering the power series expansion of the exponential, and the highest power of t appearing in this expansion reveals the size of a maximal pairwise edge-disjoint collection. □

3 Second quantization of graph processes

With tools in hand, graph processes can now be formulated as sequences in the algebraic probability space (\mathcal{G}, τ) . Associated with any graph process $(G_t)_{t=0}^\Omega$ is a corresponding sequence of idempotent-adjacency matrices, $(a_t)_{t=0}^\Omega$. This sequence is a quantum stochastic process.

For each $0 \leq k \leq \Omega$, define the indicator function $\chi_k : (\mathcal{G}, \tau) \rightarrow \{0, 1\}$ by

$$\chi_k(a) = \begin{cases} 1 & \text{if } \lambda(\tau(a_k^{n^2-n})) > \frac{n}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Defining the set S of all quantum stochastic processes associated with graph processes on n vertices, (S, τ) is an algebraic probability space.

Lemma 3.1 *On the space of quantum stochastic processes associated with graph processes on $n > 1$ vertices, define the random variable*

$$X(\omega) = \sum_{k=1}^{\infty} 2^{-k} \chi_k(\omega). \tag{3.2}$$

Then, the time step k at which a giant component first emerges in the corresponding graph sequence is given by

$$k_0 = 1 - \log_2 X(\omega). \tag{3.3}$$

Proof Begin by noting that the value k_0 corresponds to the first value of k for which $\chi_k(\omega) = 1$. It then follows from the identity

$$1 = \sum_{k=1}^{\infty} 2^{-k} = \sum_{k=1}^{k_0-1} 2^{-k} + \sum_{k_0}^{\infty} 2^{-k} \tag{3.4}$$

that

$$1 - \sum_{k=1}^{k_0-1} 2^{-k} = 2^{-k_0+1}. \tag{3.5}$$

Hence, $k_0 = -\log_2 X(\omega) + 1$. □

The method of second quantization refers to the extension of operator-theoretic models of single-particle systems to systems of arbitrarily many particles. In quantum probability theory, a single particle can be represented in a Hilbert space \mathcal{H} . In order to work with a system of arbitrarily many particles, an infinite-dimensional Hilbert space is constructed. For example, the Hilbert space $\bigoplus_{n=1}^{\infty} \mathcal{H}^{\otimes n}$ is referred to as the *free Fock space over \mathcal{H}* . The n th direct summand is the *n -particle subspace* (cf. [18]).

Second quantization associates the Hilbert space with the corresponding Fock space. Operators on the finite-dimensional subspaces are extended to operators on the Fock space.

Borrowing the notion of creation operators from quantum probability, we can think of graphs on n vertices as systems of some number of particles between 0 and Ω . At each step of the process, an edge is “created” between a randomly chosen pair of non-adjacent vertices. Hence, the N th graph of the process $(G_t)_{t=0}^{\Omega}$ will correspond to an N -particle system. This system can be in any one of $\binom{\Omega}{N}$ states, depending on which edges are present.

The goal now is to create a single operator that encodes all possible graph processes on n vertices. For fixed $n > 0$, consider the vertex set $V = \{1, 2, \dots, n\}$. For each $1 \leq i \leq \Omega$, let a_i denote the idempotent-adjacency matrix associated with $G = (V, E)$ where $|E| = 1$. In other words, the collection $\{a_i\}$ represents all idempotent-adjacency matrices of one-edge subgraphs of the complete graph K_n .

Define

$$\Gamma_1 = a_1 \otimes a_2 \otimes \cdots \otimes a_\Omega. \tag{3.6}$$

By construction, Γ_1 encodes all one-step graph processes on n vertices. Extending this idea to N -step graph processes, define the operator Γ_N by

$$\Gamma_N := \bigotimes_{i_1=1}^{\Omega} \bigotimes_{i_2=i_1+1}^{\Omega} \cdots \bigotimes_{i_N=i_{N-1}+1}^{\Omega} (a_{i_1} + \cdots + a_{i_N}). \tag{3.7}$$

The operator Γ_N can now be written in the form

$$\Gamma_N = \bigotimes_{\ell=1}^{\binom{\Omega}{N}} M_\ell, \tag{3.8}$$

where each M_ℓ is the idempotent-adjacency matrix of a simple graph on n vertices having N edges. In particular, each M_ℓ represents the N th step of a graph process.

Define the *scalar sum* functional $\langle\langle \cdot \rangle\rangle : \mathcal{I}_V \otimes \mathcal{I}_{V \times V} \rightarrow \mathbb{R}$ by

$$\langle\langle x \rangle\rangle = \left\langle\left\langle \sum_{\underline{i} \in 2^V, \underline{j} \in 2^{V \times V}} \alpha_{\underline{i}, \underline{j}} \varepsilon_{\underline{i}} \otimes \gamma_{\underline{j}} \right\rangle\right\rangle = \sum_{\underline{i} \in 2^V, \underline{j} \in 2^{V \times V}} \alpha_{\underline{i}, \underline{j}}. \tag{3.9}$$

Finally, define the mapping τ^* by

$$\tau^*(M_1 \otimes \cdots \otimes M_{\binom{\Omega}{N}}) = \prod_{\ell=1}^{\binom{\Omega}{N}} \exp(\lambda(\tau(M_\ell))). \tag{3.10}$$

Proposition 3.2 *Let a graph process (G_t) be given. Let X_N denote the size of a maximal connected component of G_N . Then, the expected value of X_N is given by*

$$\mathbb{E}(X_N) = \frac{N!(\Omega - N)!}{\Omega!} \ln(\tau^*((\Gamma_N)^{n^2-n})). \tag{3.11}$$

Proof By construction, $\ln(\tau^*((\Gamma_N)^{n^2-n}))$ is the sum of maximal component sizes taken over all graphs occurring in the N th step of the process. There are $\binom{\Omega}{N}$ such graphs, and all occur with equal probability. Hence, the result. \square

Define the mapping $\nu_\kappa : \mathcal{G} \rightarrow \{0, 1\}$ by

$$\nu_\kappa(M) = \begin{cases} 1 & \text{if } \lambda(\tau(M)) = \kappa, \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

Now, define ν_κ^* by

$$\nu_\kappa^*(M_1 \otimes \cdots \otimes M_{\binom{\Omega}{N}}) = \prod_{\ell=1}^{\binom{\Omega}{N}} \exp(\nu_\kappa(M_\ell)). \tag{3.13}$$

Proposition 3.3 *Let a graph process (G_t) be given. Let $M \leq n$ be an arbitrary positive integer. Let $E_{N,\kappa}$ be the event that G_N contains a maximal connected component of size κ . Then,*

$$\mathbb{P}(E_{N,\kappa}) = \frac{N!(\Omega - N)!}{\Omega!} \ln(v_\kappa^*((\Gamma_N)^{n^2-n})). \tag{3.14}$$

Proof As in the proof of Proposition 3.2, each graph occurs with equal probability $\binom{\Omega}{N}$, and $\ln(v_\kappa^*((\Gamma_N)^{n^2-n}))$ represents the number of N -edge graphs containing a maximal component of size κ . □

The following corollary is an immediate consequence of the preceding results using simple inclusion–exclusion.

Corollary 3.4 *Let a graph process (G_t) be given. Let $\kappa \leq n$ be an arbitrary positive integer. Let $X_{N,\kappa}$ denote the event that a maximal connected component of size κ emerges at time step N . Then,*

$$\begin{aligned} \mathbb{P}(X_{N,\kappa}) \leq & \frac{(N - 1)!(\Omega - (N - 1))!}{\Omega!} (\ln(\mu^*(\Gamma_{N-1})) - \ln(v_\kappa^*((\Gamma_{N-1})^{n^2-n}))) \\ & + \frac{N!(\Omega - N)!}{\Omega!} \ln(v_\kappa^*((\Gamma_N)^{n^2-n})). \end{aligned} \tag{3.15}$$

By considering a second quantization using vertex/edge-labeled idempotent-adjacency matrices, it becomes possible to compute the expected number of (k, d) -components of the N th graph of the process. In particular, the expected number of spanning trees of G_N can be computed.

By considering vertex/edge-labeled idempotent-adjacency matrices $\{\hat{a}_i\}$ in place of the matrices $\{a_i\}$ used to construct Γ_N , the second quantization operator Υ_N is analogously defined.

That is,

$$\Upsilon_N := \bigotimes_{i_1=1}^{\Omega} \bigotimes_{i_2=i_1+1}^{\Omega} \cdots \bigotimes_{i_N=i_{N-1}+1}^{\Omega} (\hat{a}_{i_1} + \cdots + \hat{a}_{i_N}). \tag{3.16}$$

The operator Υ_N can now be written in the form

$$\Upsilon_N = \bigotimes_{\ell=1}^{\binom{\Omega}{N}} \hat{M}_\ell, \tag{3.17}$$

where each \hat{M}_ℓ is the vertex/edge-labeled idempotent-adjacency matrix of a simple on n vertices having N edges; i.e., simple graphs representing N th steps of graph processes.

Define the mapping $d^* : \mathcal{G}_e^{\otimes \binom{\Omega}{N}} \rightarrow \mathbb{R}$ by

$$d^*(\hat{M}_1 \otimes \cdots \otimes \hat{M}_{\binom{\Omega}{N}}) = \prod_{\ell=1}^{\binom{\Omega}{N}} \exp(\dim(\langle \tau(\hat{M}_\ell^{n^2-n}) \rangle_{n,n-1})). \tag{3.18}$$

The following proposition follows from Proposition 2.14 and the construction of the second quantization operator.

Proposition 3.5 *Let a graph process (G_t) be given. Let T_N denote the number of spanning trees of G_N . Then, the expected value of T_N is given by*

$$\mathbb{E}(T_N) = \frac{N!(\Omega - N)!}{\Omega!} \ln(d^*(\hat{Y}_N)). \tag{3.19}$$

Proof By definition,

$$\mathbb{E}(T_N) = \sum_{k \geq 0} k \mathbb{P}(T_N = k). \tag{3.20}$$

Since the graphs G occurring in the N th step of the process are mutually exclusive,

$$\begin{aligned} \sum_k k \mathbb{P}(T_N = k) &= \sum_k k \sum_{\substack{G \text{ having } k \\ \text{spanning trees}}} \mathbb{P}(G_N = G) \\ &= \sum_k k \cdot \#\{N\text{th Graphs with } k \text{ spanning trees}\} \cdot \frac{N!(\Omega - N)!}{\Omega!} \\ &= \frac{N!(\Omega - N)!}{\Omega!} \sum_{N\text{-edge graphs } G_N} \#\{\text{spanning trees in } G_N\} \\ &= \frac{N!(\Omega - N)!}{\Omega!} \ln(d^*(\hat{Y}_N)). \end{aligned} \tag{3.21}$$

□

4 Conclusion

This paper represents one step toward a comprehensive study of graph processes and algorithms using tools of algebraic probability.

References

1. Accardi, L., Ben Ghorbal, A., Obata, N.: Monotone independence, comb graphs and Bose–Einstein condensation. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **7**, 419–435 (2004)
2. Berezin, F.A.: *The Method of Second Quantization*. Academic Press, New York (1966)
3. Bohman, T., Frieze, A.: Avoiding a giant component. *Random Struct. Algorithms* **19**, 75–85 (2001)

4. Bohman, T., Frieze, A., Wormald, N.: Avoiding a giant component in half the edge set of a random graph. *Random Struct. Algorithms* **25**, 432–449 (2004)
5. Bohman, T., Kim, J.H.: A phase transition for avoiding a giant component. *Random Struct. Algorithms* **28**, 195–214 (2006)
6. Bohman, T., Kravitz, D.: Creating a giant component. *Comb. Probab. Comput.* **15**, 489–511 (2006)
7. Bollobás, B.: The evolution of random graphs. *Trans. Am. Math. Soc.* **286**, 257–274 (1984)
8. Chung, F., Lu, L.: Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6**, 125–145 (2002)
9. Erdős, P., Rényi, A.: On random graphs. I. *Publ. Math. (Debr.)* **6**, 290–297 (1959)
10. Erdős, P., Rényi, A.: On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.* **5**, 17–61 (1960)
11. Erdős, P., Rényi, A.: On the evolution of random graphs. *Bull. Inst. Int. Stat. Tokyo* **38**, 343–347 (1961)
12. Flaxman, A., Gamarnik, D., Sorkin, G.: Embracing the giant component. *Random Struct. Algorithms* **27**, 277–289 (2005)
13. Frieze, A.M., Łuczak, T.: Edge-disjoint spanning trees in random graphs. *Period. Math. Hung.* **21**, 35–37 (1990)
14. Hashimoto, Y., Hora, A., Obata, N.: Central limit theorems for large graphs: Method of quantum decomposition. *J. Math. Phys.* **44**, 71–88 (2003)
15. Molloy, M.: When does the giant component bring unsatisfiability? *Combinatorica* **28**, 693–734 (2008)
16. Obata, N.: Quantum probabilistic approach to spectral analysis of star graphs. *Interdiscip. Inf. Sci.* **10**, 41–52 (2004)
17. Palmer, E.M., Spencer, J.J.: Hitting time for k edge-disjoint spanning trees in a random graph. *Period. Math. Hung.* **31**, 235–240 (1995)
18. Parthasarathy, K.R.: *An Introduction to Quantum Stochastic Calculus*. Birkhäuser, Basel (1992)
19. Schott, R., Staples, G.S.: Nilpotent adjacency matrices and random graphs. *Ars Comb.* **98**, 225–239 (2011)
20. Schott, R., Staples, G.S.: Nilpotent adjacency matrices, random graphs, and quantum random variables. *J. Phys. A, Math. Theor.* **41**, 155205 (2008)
21. Staples, G.S.: Clifford-algebraic random walks on the hypercube. *Adv. Appl. Clifford Algebras* **15**, 213–232 (2005)